# RL+Robot

## 以下分别是参考的各类论文

最初了解火灾等内容：

第39卷第9期
2022年9月

计 算 机 应 用 研 究
Application Research of Computers

Vol. 39 No. 9
Sep. 2022

### 基于FDS和元胞自动机动态耦合的火灾疏散模型*

李　超[a,b]，李宇飞[a,b]，霍非舟[a,b†]，张钦钦[a,b]
（武汉理工大学 a. 中国应急管理研究中心；b. 安全科学与应急管理学院，武汉 430070）

**摘　要：**为研究火灾场景下温度、烟气和CO浓度等灾害因子对疏散的影响，建立基于FDS和元胞自动机动态耦合的火灾疏散模型。将FDS的网格和元胞自动机的元胞一一对应，将由FDS运行得到的灾害数据通过Python等技术手段实时加载到元胞中，使灾害数据持续影响行人转移概率，从而实现灾害和疏散的动态耦合。以单层教学楼作为仿真场景进行模拟分析，对火源位置和热释放速率等因素进行讨论，得出这些因素对行人疏散进程的影响规律；将模型与传统软件和同类方案进行对比。研究表明，火灾导致的高温和烟气会影响行人对疏散路径和安全出口的选择；热释放速率越大，行人越早处于危险状态，同时处于危险状态的行人也越多。该模型相比传统疏散软件不仅能考虑火灾产生的致灾因子对行人疏散的动态影响，还能确定行人最早处于危险状态的位置和时间，并用可视化的方式表现出来。

**关键词：**火灾疏散；FDS；元胞自动机；动态耦合；仿真模拟
**中图分类号：**TP391.9　　**文献标志码：**A　　**文章编号：**1001-3695(2022)09-032-2768-05
doi:10.19734/j.issn.1001-3695.2022.03.0086

### Fire evacuation model based on dynamic coupling of FDS and cellular automata

Li Chao[a,b], Li Yufei[a,b], Huo Feizhou[a,b†], Zhang Qinqin[a,b]
(*a. China Research Center for Emergency Management, b. School of Safety Science & Emergency Management, Wuhan University of Technology, Wuhan* 430070, *China*)

**Abstract:** To study the influence of disaster factors such as temperature, smoke, and CO concentration on evacuation at the fire scene, this paper established a fire evacuation model based on the dynamic coupling of FDS and cellular automata. It mapped the grid of FDS and the cell of cellular automata one by one, and loaded the disaster data obtained from the operation of FDS into the cell in real-time through Python and other technical means, so that the disaster data could continuously affect the pedestrian transfer probability, to realize the dynamic coupling between disaster and evacuation. Taking the single-story teaching building as the simulation scene for simulation analysis, this paper discussed the factors such as fire source location and heat release rate, and obtained the influence law of these factors on pedestrian evacuation process. Compared the model with traditional software and similar schemes, the result shows that the high temperature and smoke caused by fire will affect the choice of pedestrian evacuation path and emergency exit, the greater the heat release rate, the earlier pedestrians are in danger, and the more pedestrians are in danger at the same time. Compared with the traditional evacuation software, the model can not only consider the dynamic impact of the disaster causing factors caused by fire on pedestrian evacuation, but also determine the location and time of pedestrians in the first dangerous state, which can be displayed visually.

**Key words:** fire evacuation; FDS; cellular automata; dynamic coupling; analogue simulation

#### 0　引言

近年来，建筑火灾频发，造成了大量的人员伤亡和财产损失。据近10年的数据统计，我国一共发生过3.1万起高层建筑火灾，死亡人数474人，直接经济损失约15.6亿元[1]。发生火灾后，人员在恐慌、从众、冲动等心理特征下，难以冷静应对形势并作出理性的逃离决策。因此，为最大程度地减少火灾造成的人员伤亡，开展火灾条件下的人员疏散动力学研究是非常必要的。考虑到火灾条件的危险性，在现实中不可能进行火灾疏散实验，随着计算机技术的发展，计算机模拟成为一个非常重要而可行的工具。国际上开发的较为成熟的行人疏散模拟软件包括Pathfinder[2]、Building EXODUS[3]和FDS + EVAC[4]

建筑发生火灾后的可用疏散时间(ASET)，然后再通过Pathfinder等软件计算出该建筑行人的必须疏散时间(RSET)，最后通过比较RSET与ASET的大小来确定建筑是否满足防火和疏散的要求，这样就不能很好地考虑到火灾灾害因子对疏散过程的影响[1,5]。

鉴于上述原因，越来越多的研究人员将目光投放到建立疏散模型上。总体上讲，疏散模型可大致分为宏观模型和微观模型两类[6]，宏观模型将人群移动视为流体运动，能够高效计算出大规模人群的疏散时间，但宏观模型过于理想化，无法反映个体之间的相互作用和异质性；作为微观模型典型代表的元胞自动机模型不仅能够反映出行人个体间的差异还能体现出行人在疏散过程中典型的心理特征和行为反应，受到广大研究者

等，在工程实践中具有较高的应用价值。但仿真软件也基于固定的行人疏散模型，用户不能改变疏散的实际移动规则，这是疏散软件最大的局限。例如，一些学者通过使用 FDS 计算出

的关注。Zheng 等人[7-9]考虑了火灾和烟气对行人运动的影响以及烟层场对疏散过程的影响，改进了基于场域模型的元胞自动机模型；金泽人等人[10]考虑了火灾导致的恐慌心理对行

收稿日期：2022-03-08；修回日期：2022-04-26　基金项目：国家自然科学基金资助项目(51808422)；中央高校基本科研业务费专项资金资助项目(2021 Ⅲ052JC)

作者简介：李超(1998-)，男，山西大同人，硕士研究生，主要研究方向为计算机仿真、建筑性能化防火设计、智慧城市安全、行人疏散等；李宇飞(1997-)，男，湖北天门人，硕士研究生，主要研究方向为行人疏散动力学；霍非舟(1987-)，男(通信作者)，河南周口人，副教授，博士，主要研究方向为人员疏散、风险评估、公共安全与应急管理(huofz@whut.edu.cn)；张钦钦(1999-)，女，山东枣庄人，硕士研究生，主要研究方向为行人疏散.

基于ＦＤＳ和元胞自动机动态耦合的火灾疏.pdf

健康值等参照来源：

环境：多层建筑火灾场景（详细）

奖励：集体奖励（人员疏散和死亡以及机器人成功引导人类）个体（每秒该区域人越多负奖励越多，机器人遇火，人遇火）详细。

观测：所有人类坐标，机器人坐标，自疏散开始的时间，集体健康值。

动作：机器人动作决策由两个浮点数组成，计划下一个水平位置坐标

人员移动：unity导航

# Adversarial Reinforcement Learning for Enhanced Decision-Making of Evacuation Guidance Robots in Intelligent Fire Scenarios

Hantao Zhao, Zhihao Liang, Tianxing Ma, Xiaomeng Shi, Mubbasir Kapadia, Tyler Thrash, Christoph Hoelscher, Jinyuan Jia, Bo Liu, and Jiuxin Cao

***Abstract*—In the context of rapid urbanization, traditional manual guidance and static evacuation signs are increasingly inadequate for addressing complex and dynamic emergencies. This study proposes an innovative emergency evacuation framework that optimizes the crowd evacuation by integrating multiagent reinforcement learning (MARL) with adversarial reinforcement learning (ARL). The developed simulation environment models realistic human behavior in complex buildings and incorporates robotic navigation and intelligent path planning. A novel simulated human behavior model was integrated, capable of complex human–robot interaction, independent escape route searching, and exhibiting herd mentality and memory mechanisms. We also proposed a multiagent framework that combines MARL and ARL to enhance overall evacuation efficiency and robustness. Additionally, we developed a new ARL evaluation framework that provides a novel method for quantifying agents' performance. Various experiments of differing difficulty levels were conducted, and the results demonstrate that the proposed framework exhibits advantages in emergency evacuation scenarios. Specifically, our ARLR approach increased survival rates by 1.8% points in low-difficulty evacuation tasks compared to the RLR approach using only MARL algorithms. In high-difficulty evacuation tasks, the ARLR approach raised survival rates from 46.7% without robots to 64.4%, exceeding the RLR approach by 1.7% points. This study aims to enhance the efficiency and safety of human–robot collaborative fire evacuations and provides theoretical support for evaluating and improving the performance and robustness of ARL agents.**

***Index Terms*—Adversarial reinforcement learning (ARL), human–robot interaction, multiagent reinforcement learning (MARL), simulation frameworks.**

## I. Introduction

THE accelerated urbanization process has led to a significant increase in the density and complexity of public buildings, posing severe challenges to the effectiveness of traditional evacuation methods during emergencies like fires, such as manual guidance and static signs [1], [2]. This, in turn, constitutes a major threat to public safety. In response, researchers are actively developing new technologies and methods that integrate human behavior models [3], simulation tools [4], [5], and optimization algorithms [6], [7]—such as adaptive signage systems and evacuation-assistance robots [8]—to enhance evacuation efficiency and safety. These innovative approaches aim to meet the urgent evacuation demands in complex environments and provide scientific support for improving public safety levels [9], [10].

Accurate guide information and the visibility of emergency evacuation signs are key factors for optimizing crowd evacuation efficiency [11], [12]. Studies have shown that dynamic signage systems [13], [14] and intelligent interactive human–robot collaborative evacuation systems [15], [16], such as evacuation robots utilizing sensors and map data, can effectively guide crowds to safety during emergencies [17], [18]. Moreover, the behavior of robots within the crowd has been proven to significantly enhance overall evacuation performance [19].

Despite some progress, numerous challenges remain in practical application, such as ethical constraints and

Received 29 April 2024; revised 22 August 2024 and 10 October 2024; accepted 12 November 2024. This work was supported in part by the Ministry of Education in China (MOE) Project of Humanities and Social Sciences under Project 24YJCZH436 and in part by the National Natural Science Foundation of China under Grant 62472092, Grant 62172089, Grant 5243000365, and Grant 71971056. *(Corresponding author: Hantao Zhao.)*

Hantao Zhao and Jiuxin Cao are with the School of Cyber Science and Engineering, Southeast University, Nanjing 211189, China, and also with Purple Mountain Laboratories, Nanjing 211111, China (e-mail: htzhao@seu.edu.cn; jx.cao@seu.edu.cn).

Zhihao Liang is with the School of Cyber Science and Engineering, Southeast University, Nanjing 211189, China (e-mail: zhliang@seu.edu.cn).

Tianxing Ma is with the Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100085, China (e-mail: matianxing@iie.ac.cn).

Xiaomeng Shi is with the School of Transportation, Southeast University, Nanjing 211189, China (e-mail: shixiaomeng@seu.edu.cn).

Mubbasir Kapadia is with the Department of Computer Science, Rutgers University, Newark, NJ 07102 USA (e-mail: mk1353@cs.rutgers.edu).

Tyler Thrash is with the Department of Biology, Saint Louis University, St. Louis, MO 63103 USA (e-mail: tyler.thrash@slu.edu).

Christoph Hoelscher is with the Chair of Cognitive Science, ETH Zurich, 8092 Zurich, Switzerland (e-mail: choelsch@ethz.ch).

Jinyuan Jia is with The Hong Kong University of Science and Technology, Guangzhou 511453, China, and also with the Game School, Jilin Animation Institute, Changchun 130012, China (e-mail: jinyuanjia@hkust-gz.edu.cn).

Bo Liu is with Purple Mountain Laboratories, Nanjing 211111, China, and also with the School of Computer Science and Engineering, Southeast University, Nanjing 211189, China (e-mail: bliu@seu.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TCSS.2024.3502420, provided by the authors.

Digital Object Identifier 10.1109/TCSS.2024.3502420



Adversarial Reinforcement Learning for Enhanced Decision-Making of Evacuation Guidance Robots i…

Chinese主要参照的环境等内容：

环境：长12米、宽10米的典型单出口场景。（√）

奖励：时间步长和总疏散时间组成。DRL控制的机器人场景的疏散时间更短，则给予较大的正奖励（？）

观测：行人位置、火源位置被编码并输入到神经网络中，用于指导机器人的行动。

动作：机器人的动作空间包含五种可能的运动决策，即“向上移动”、“向下移动”、“向左移动”、“向右移动”和“保持静止”。在每个时间步长（0.005秒）内，机器人会根据神经网络的输出选择一个动作来更新其位置。机器人的移动距离为0.005米。

Chinese Journal of Physics 92 (2024) 494–531

Contents lists available at ScienceDirect

Chinese Journal of Physics

journal homepage: www.sciencedirect.com/journal/chinese-journal-of-physics

# Robot-assisted pedestrian evacuation in fire scenarios based on deep reinforcement learning

Chuan-Yao Li [a], Fan Zhang [a], Liang Chen [b,*]

[a] School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China
[b] Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China

ARTICLE INFO

*Keywords:*
Fire evacuation
Deep reinforcement learning
Social force model
Human-robot interaction

ABSTRACT

Indoor fires pose a significant challenge to the safe evacuation of pedestrians. In response to fire hazards, pedestrians instinctively seek alternative evacuation routes to avoid the hazard zone. However, the specific location and intensity of the fire hazard zone can influence pedestrians' decisions, leading to varying congestion levels in different areas. To address this challenge and enhance overall evacuation efficiency, this paper introduces an improved social force model to depict pedestrian movement in fire scenarios and proposes a methodology that leverages dynamic robot for pedestrian evacuation, employing Deep Reinforcement Learning (DRL) and Human-Robot Interaction (HRI). The results show that in the no-robot scenario, pedestrians will detour according to the varying locations of fire hazard zones and emergency levels, resulting in congestion at different positions. In the static robot scenario, robots placed in different locations exhibit varied effects on evacuation depending on the fire hazard zones' locations and intensities. In the DRL-control robot scenario, the robot controlled by DRL and HRL can always navigate to the appropriate position to promote evacuation, regardless of the fire's location and emergency levels or the robot's initial placement. Furthermore, our findings reveal that strategically positioned robots can enhance evacuation efficiency by alleviating crowding and increasing the distance between pedestrians and fire hazard zones in most cases, thereby improving pedestrian safety. This study offers practical guidance for managing pedestrian evacuation during fire incidents and establishes a theoretical foundation for refining evacuation strategies and safety measures at fire scenes.

## 1. Introduction

Fire constitutes a pervasive and perilous indoor hazard [1]. The prevalence of fire incidents is striking, with an alarming seven million occurrences reported worldwide annually [2]. This staggering frequency underscores the gravity of the situation, as these incidents lead to tens of thousands of fatalities and hundreds of thousands of injuries each year. The ubiquity and severity of fire disasters highlight the urgent need for effective strategies to mitigate their impact and enhance overall safety measures.

In the realm of fire scenarios, the evacuation of pedestrians is a formidable challenge, marked by various intricacies and potential hazards [3]. The rapid dissemination of toxic gases during evacuations significantly endangers pedestrians' safety, exacerbating the hazards [4,5]. Additionally, the urgency of fire emergencies often triggers irrational behaviors, such as panic, impeding the orderly

* Corresponding author.
*E-mail address:* liangchen@bjut.edu.cn (L. Chen).

https://doi.org/10.1016/j.cjph.2024.09.008
Received 8 July 2024; Received in revised form 5 September 2024; Accepted 6 September 2024
Available online 14 September 2024

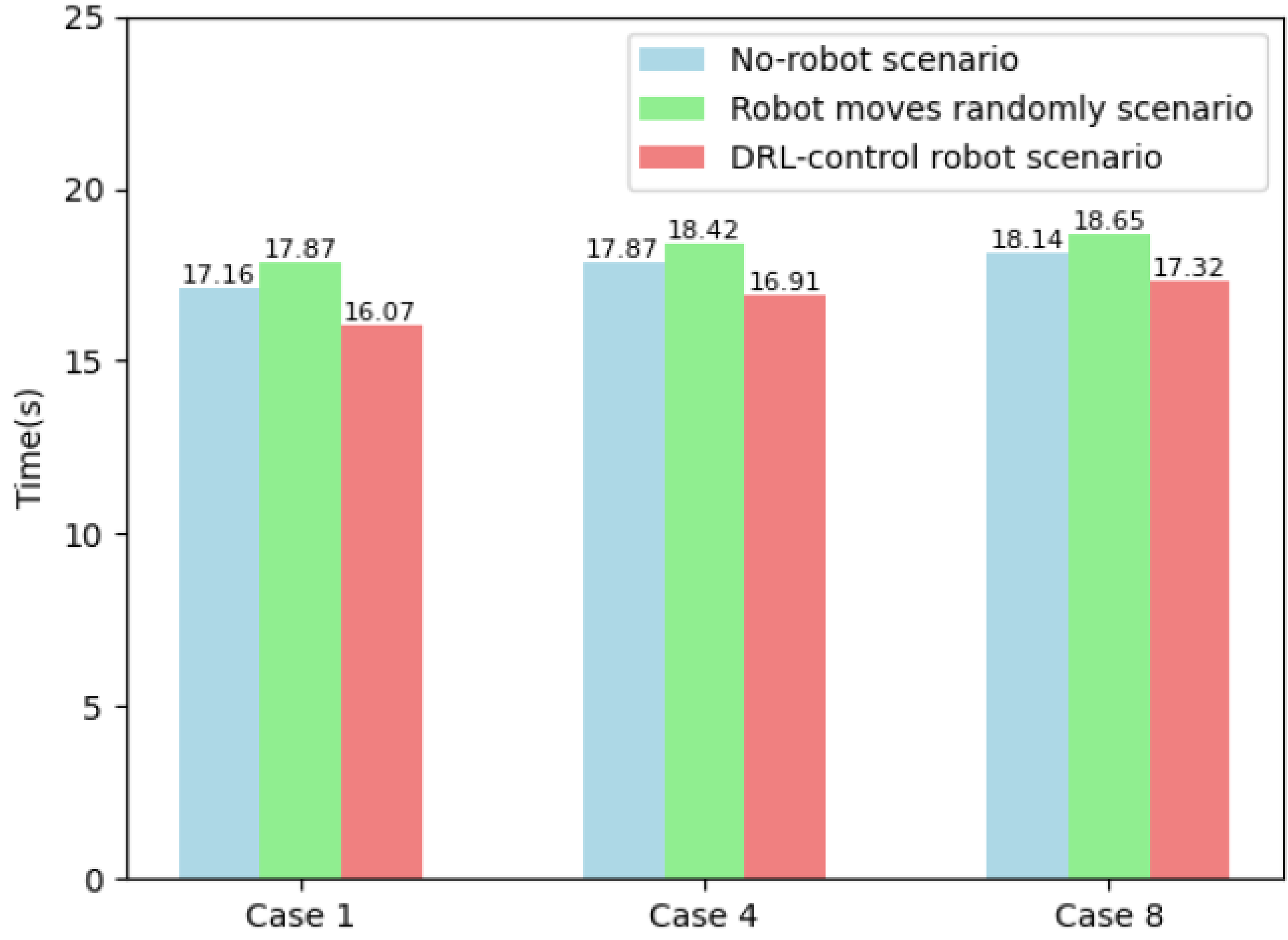



**Fig. 20.** Average evacuation time with random pedestrian distribution.

卡拉OK：

环境：卡拉OK

奖励：疏散效率，人群拥堵程度，火灾引起的危险。（√）

观测：火灾风险引起的观测，拥堵引起的观测（√）

动作：智能体可选择向前，向后，向左，向右等有限方向

Chaos, Solitons and Fractals 194 (2025) 116246

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

# Coordinating dynamic signage for evacuation guidance: A multi-agent reinforcement learning approach integrating mesoscopic crowd modeling and fire propagation

Chuan-Zhi Thomas Xie [a,b], Qihua Chen [c], Bin Zhu [a,b], Eric Wai Ming Lee [c], Tie-Qiao Tang [d], Xianfei Yin [c], Zhilu Yuan [e], Botao Zhang [c,*]

[a] School of Management, Beijing Institute of Technology, Beijing 100191, China
[b] Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing 314000, China
[c] Department of Architecture and Civil Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
[d] School of Transportation Science and Engineering, Beihang University, Beijing 100191, China
[e] State Key Laboratory of Subtropical Building and Urban Science, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

*Keywords:*
Evacuation
Dynamic signage
Crowd modeling
Fire simulation
Reinforcement learning

ABSTRACT

Recurrent fire outbreaks in indoor crowd-gathering facilities, particularly those where pedestrians are unfamiliar with the spatial layout and visibility is limited, present significant challenges to evacuation safety and efficiency. Under such conditions, traditional static signage, which directs pedestrians to the nearest exit without accounting for real-time crowd and fire dynamics, often fails to provide effective guidance. To address these limitations, under the context of dynamic signage, we propose an integrated method capable of offering coordinated, real-time directional guidance of multiple signs, i.e.,: i) the extension of a sub-region-based mesoscopic model (Cellular Transmission Model, CTM) for crowd movement simulation; ii) the adoption of PyroSim to simulate the dynamic propagation of fire and its byproducts; iii) integrating real-time simulation results from i) and ii) into a dynamic environment to optimize signage directions using the Multi-Agent Reinforcement Learning (MARL)-based QMIX algorithm, with multi-objective goals addressing evacuation efficiency, congestion levels, and fire-induced risks simultaneously. Advancements of this paper can be summarized as: i) in terms of environment construction for dynamic crowd evacuation guidance, our approach represents one of the first to integrate sub-regional mesoscopic crowd modeling with dynamic fire propagation simulation. This integration naturally aligns with the granularity of directional guidance, where pedestrians within the same sub-region receive uniform instructions; ii) regarding real-time directional guidance generation for multi-sign, our method extends the discrete MARL algorithm QMIX, which is well-suited for the discrete action space of each sign (i.e., forward, backward, left, right). This extension effectively manages the high-dimensional challenge of coordinating multiple signs simultaneously while optimizing both evacuation efficiency and safety; iii) from the perspective of model application, we demonstrate the effectiveness of our CTM-PyroSim-QMIX framework in a fire evacuation scenario in a real-world karaoke venue, characterized by low visibility and pedestrians' unfamiliarity with the layout. Benchmarking against the traditional static signage approach, we show that the directional guidance generated by our method enhances evacuation efficiency and reduces fire-related and congestion-induced hazards across 10 single and dual fire source cases. Specifically, the maximum improvements observed in evacuation efficiency, fire-related hazards, and congestion-related risks are approximately over 30 %, 50 % and 70 %, respectively.

## 1. Introduction

According to the United Nations, by 2030, it is estimated that there will be 43 megacities globally, each with a population exceeding 10 million [1]. To accommodate the diverse social and functional needs of residents, megacities are equipped with a range of working, residential,

* Corresponding author.
*E-mail address:* botzhang2-c@my.cityu.edu.hk (B. Zhang).

https://doi.org/10.1016/j.chaos.2025.116246
Received 18 November 2024; Received in revised form 21 February 2025; Accepted 27 February 2025
Available online 9 March 2025

# 实验环境等总述

## 重新梳理一下项目总成

环境：长18米、宽15米的典型单出口场景。(初始条件)

奖励：密集奖励，健康值变化，疏散：给予群体中疏散节点人员剩余健康值加50

观测：火灾风险影响的健康值，人群疏散时间

动作：智能体可在元胞里五个动作状态选择。

# 实验代码基础

## 地图的代码：

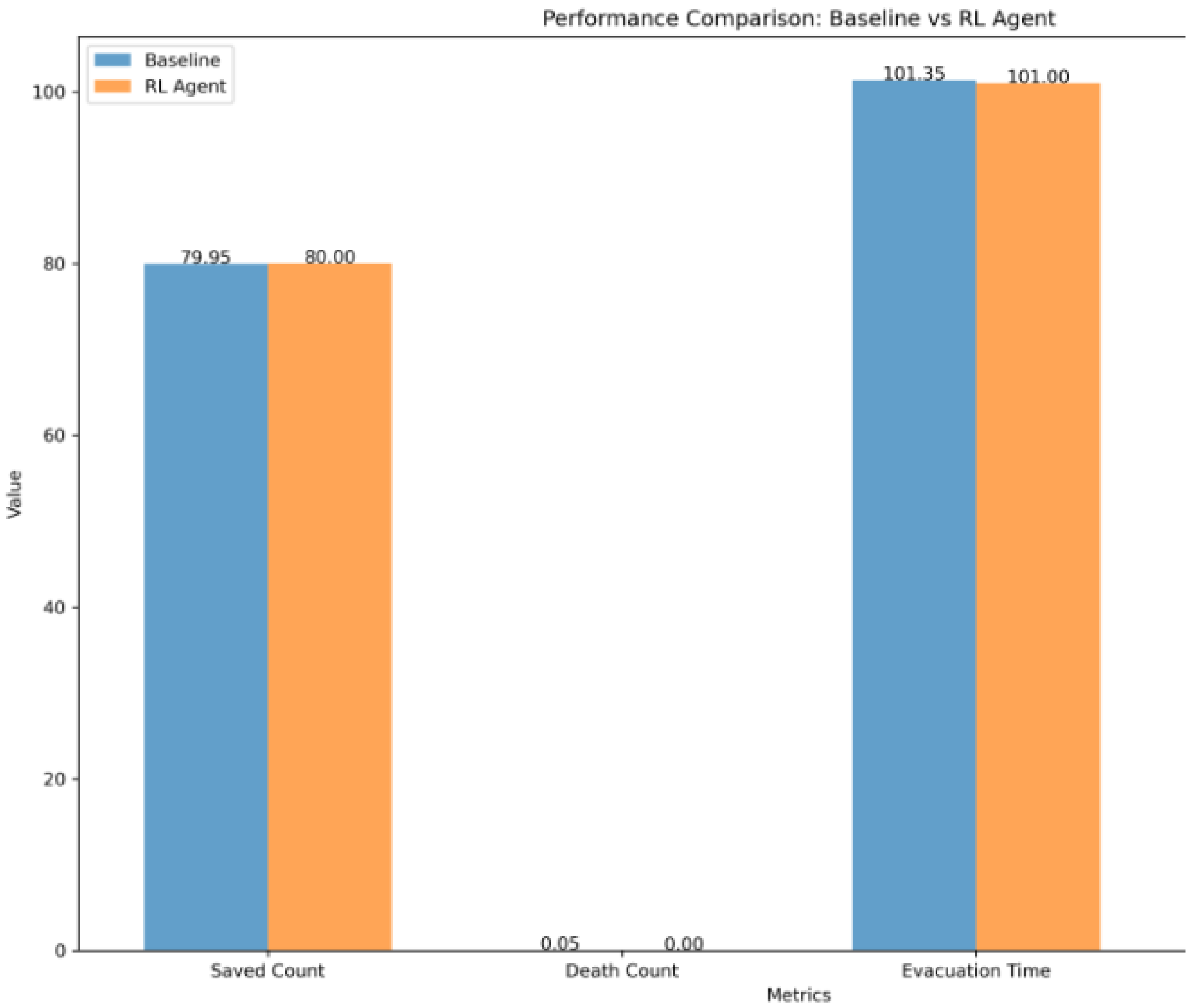
map.py

## 行人行为代码

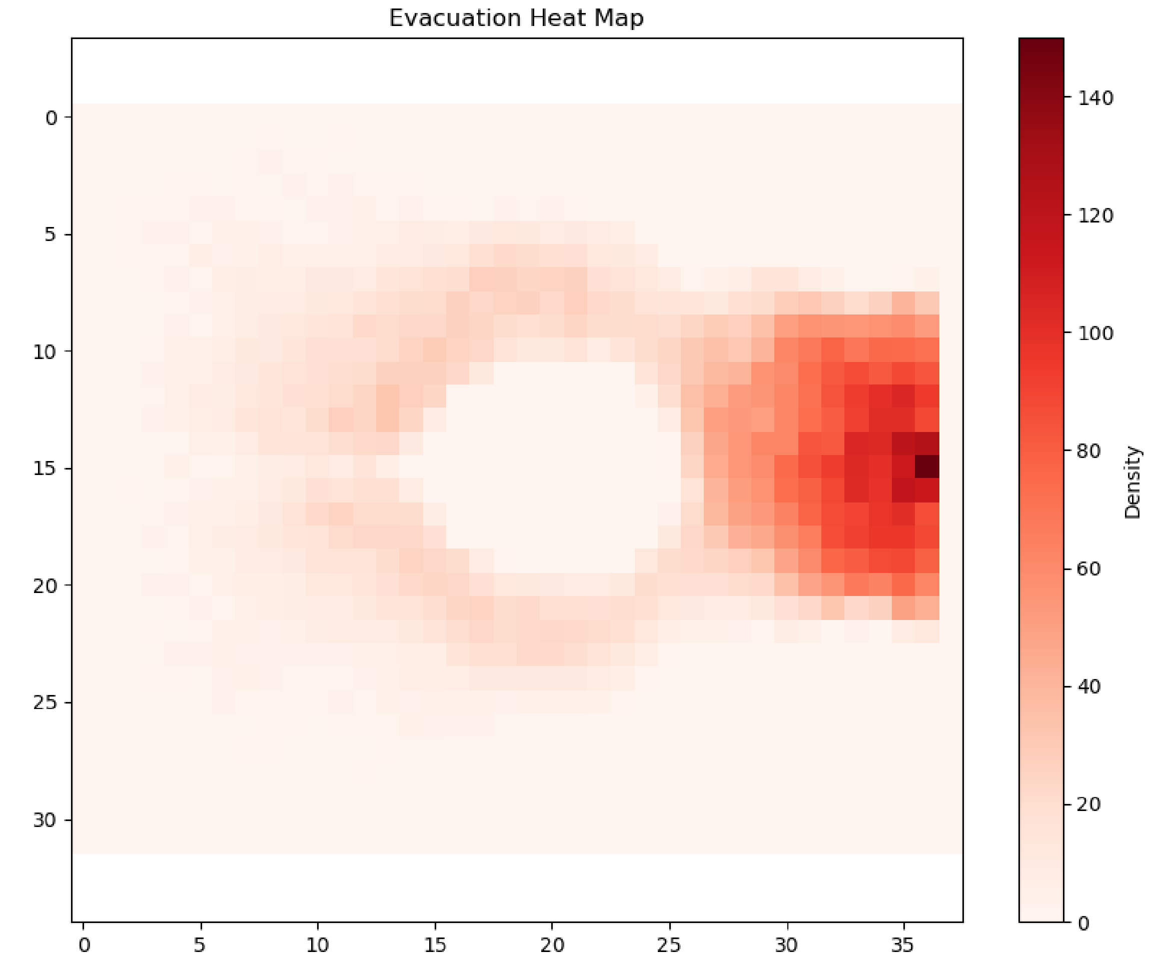
people.py

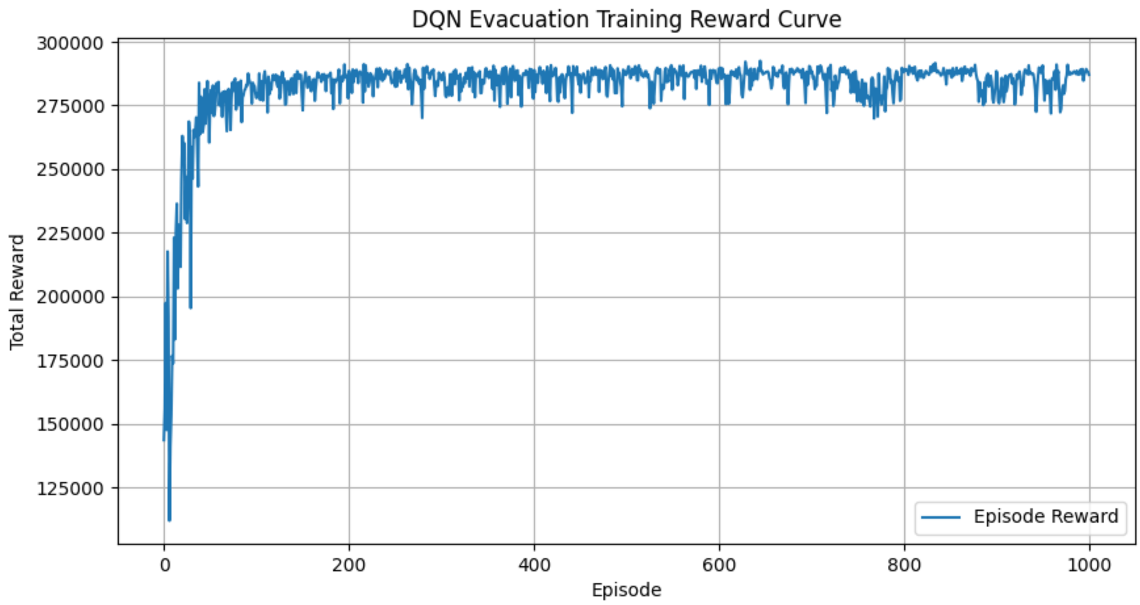
### 主函数:

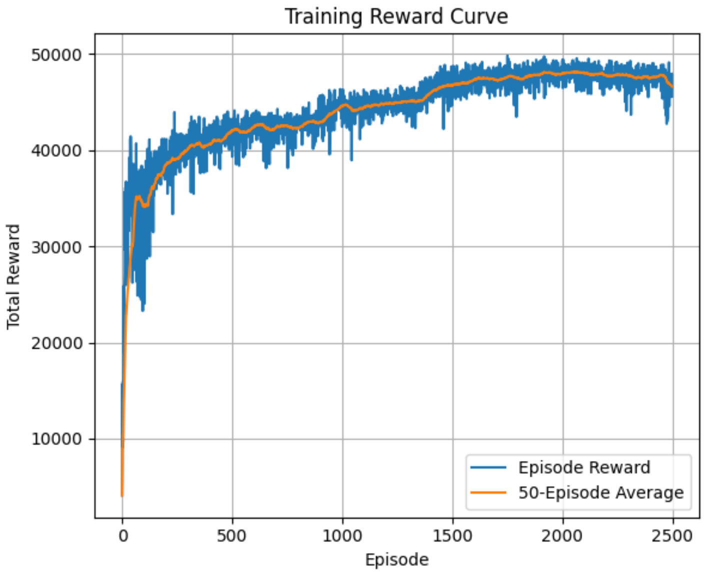
main.py


Performance Comparison: Baseline vs RL Agent
Baseline
RL Agent
Value
Metrics
Saved Count
Death Count
Evacuation Time
79.95
80.00
0.05
0.00
101.35
101.00
0
20
40
60
80
100

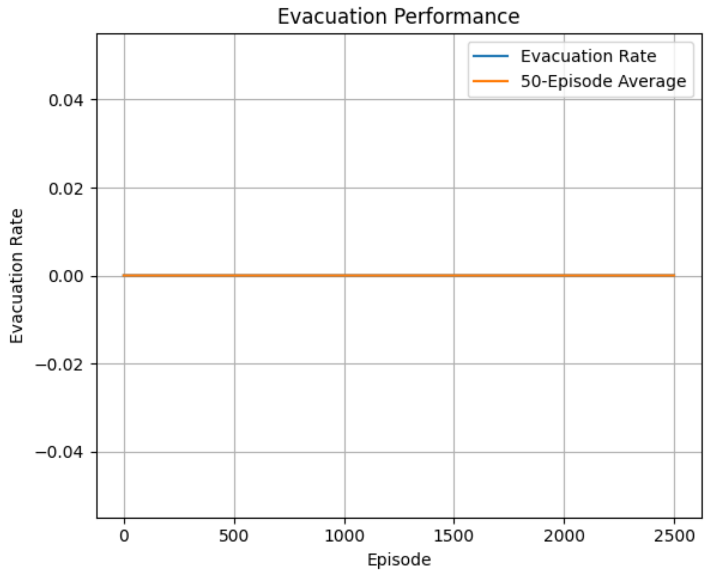
# 实验结果


Evacuation Heat Map
Density
0
5
10
15
20
25
30
35
20
40
60
80
100
120
140

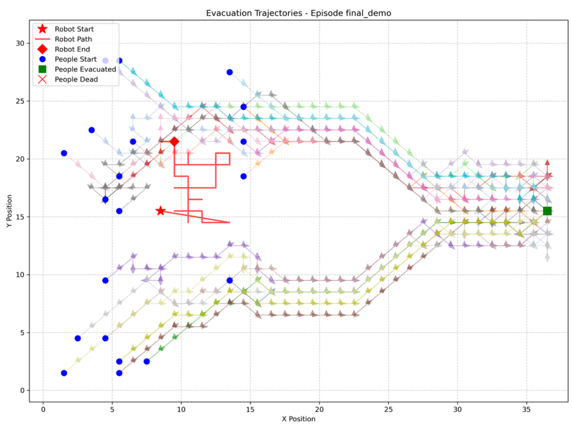

DQN Evacuation Training Reward Curve
Total Reward
Episode
Episode Reward
300000
275000
250000
225000
200000
175000
150000
125000
0
200
400
600
800
1000



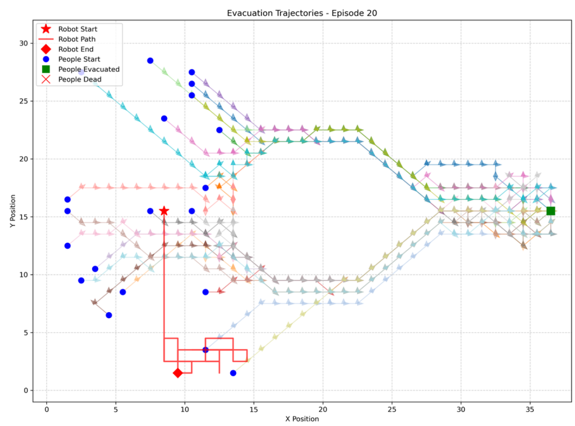
Training Reward Curve
Total Reward
Episode
Episode Reward
50-Episode Average

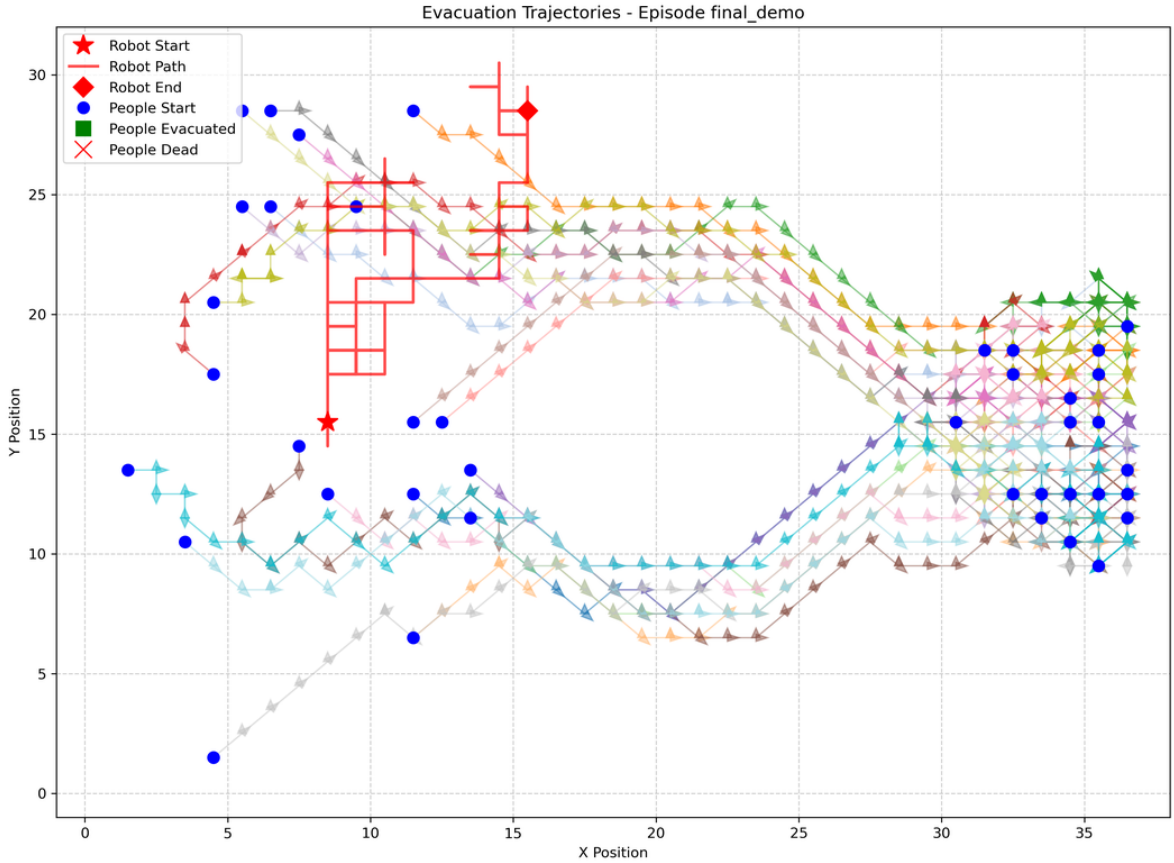
Evacuation Performance
Evacuation Rate
Episode
Evacuation Rate
50-Episode Average

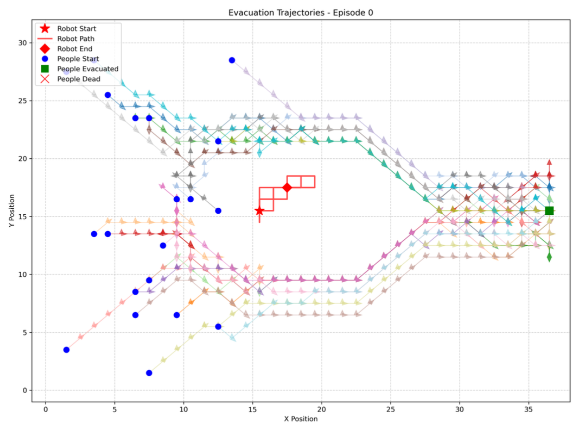



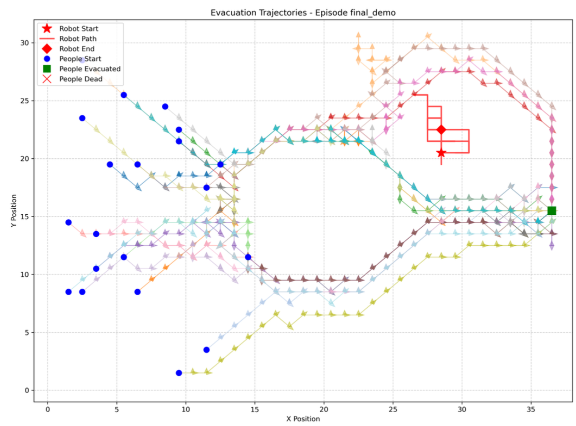
Evacuation Trajectories - Episode final_demo
Robot Start
Robot Path
Robot End
People Start
People Evacuated
People Dead
X Position
Y Position

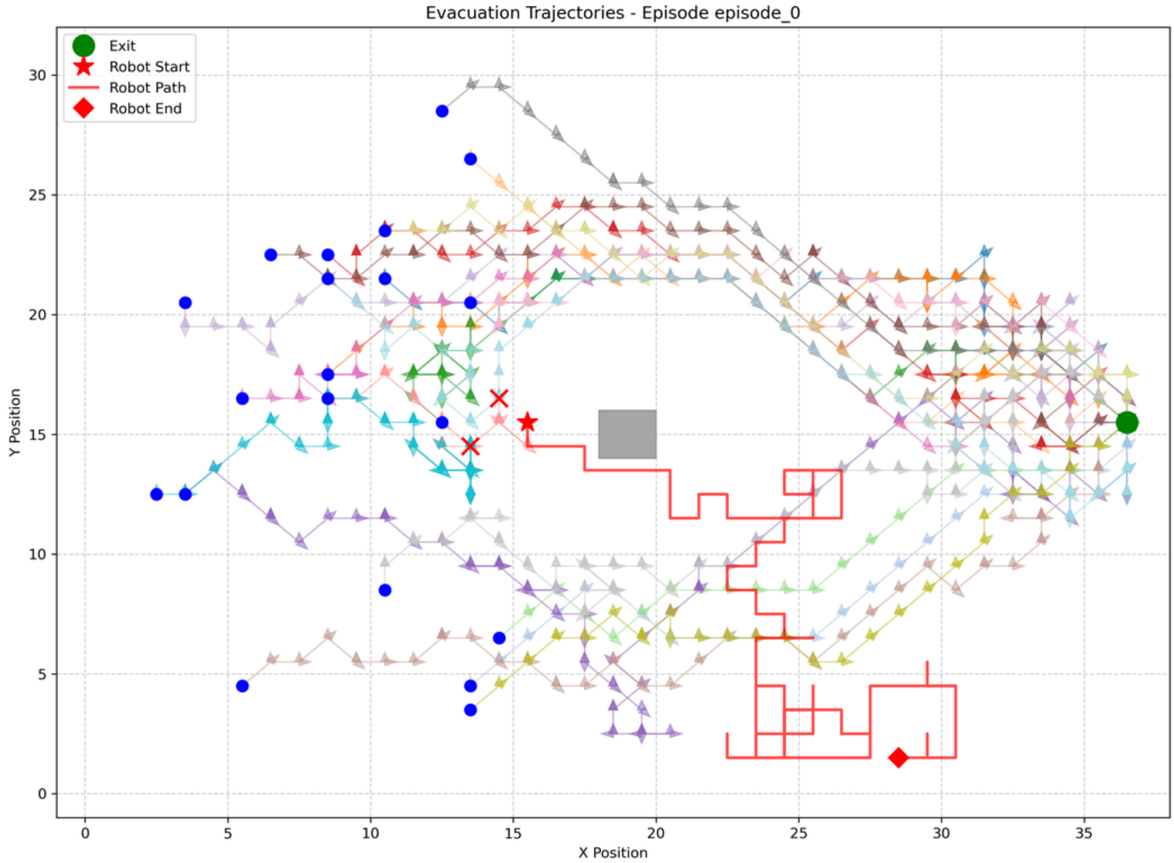
Evacuation Trajectories - Episode 20
Robot Start
Robot Path
Robot End
People Start
People Evacuated
People Dead
X Position
Y Position

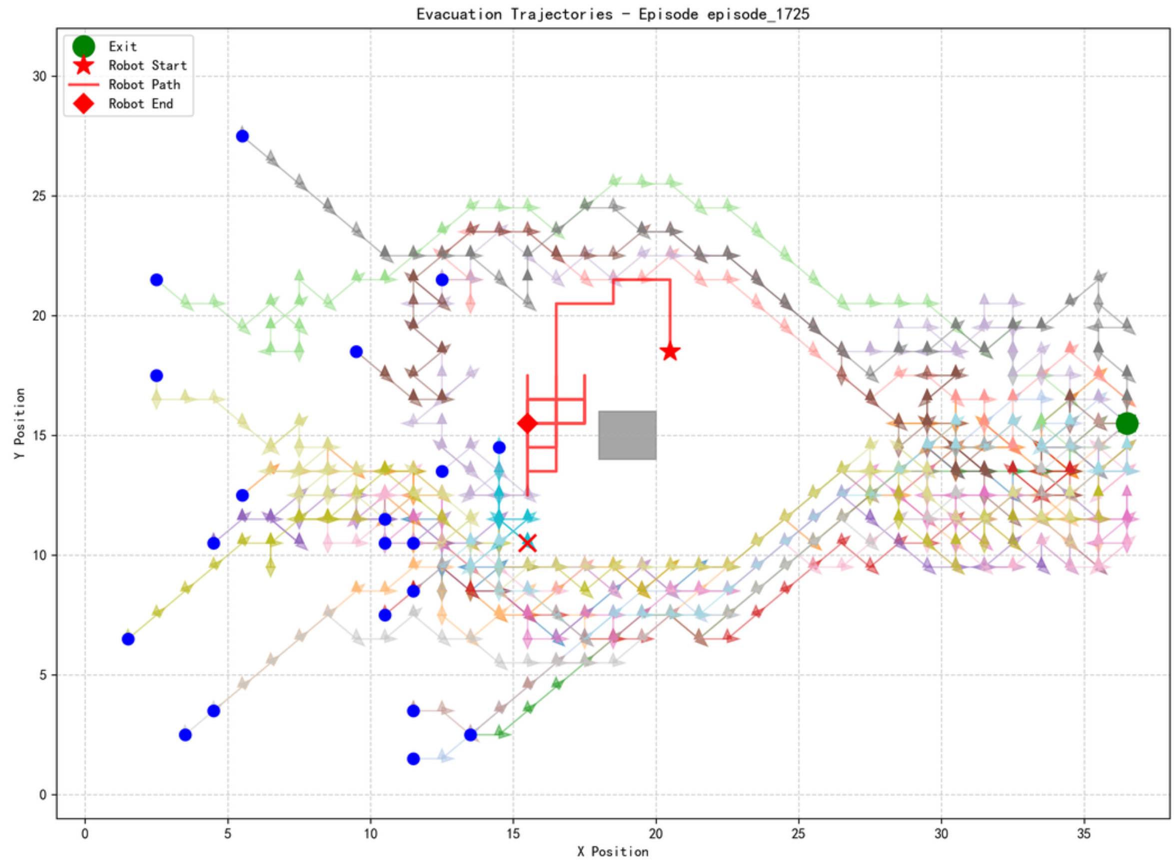


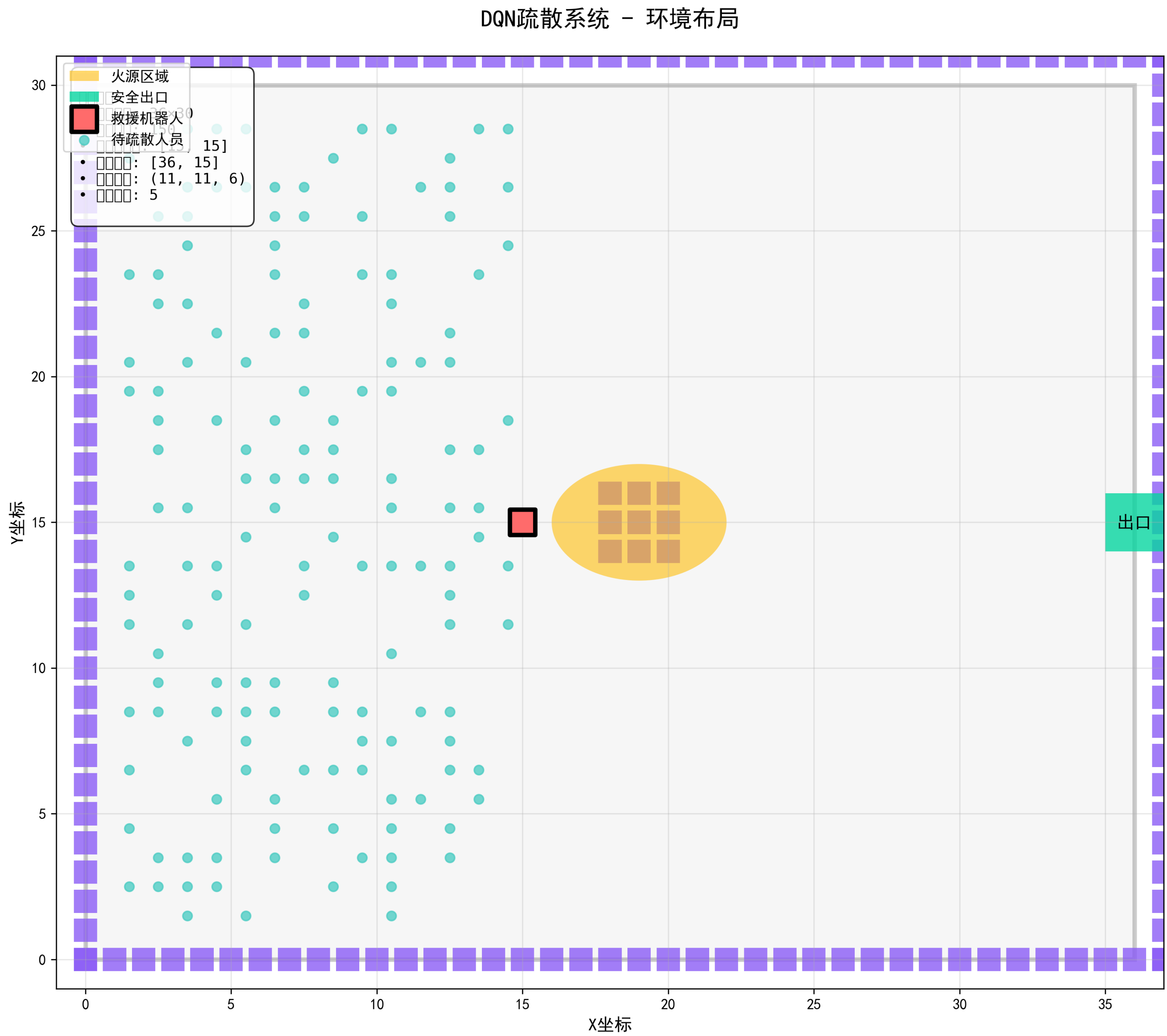

机器人会作为力影响这些人，使得人的速度变化。

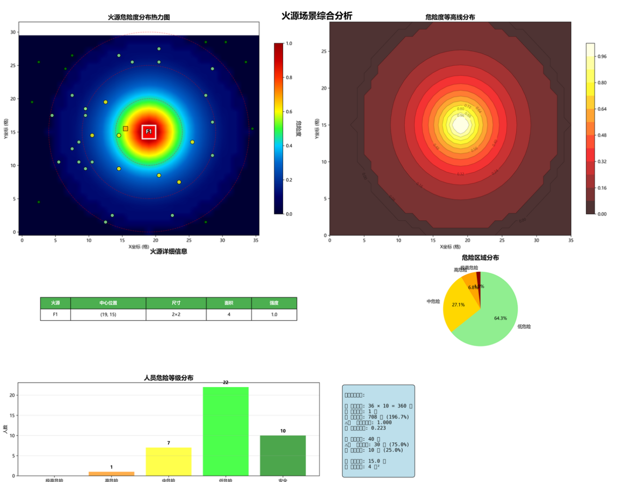

Evacuation Trajectories - Episode episode_0
Exit
Robot Start
Robot Path
Robot End
Y Position
X Position
0
5
10
15
20
25
30
35

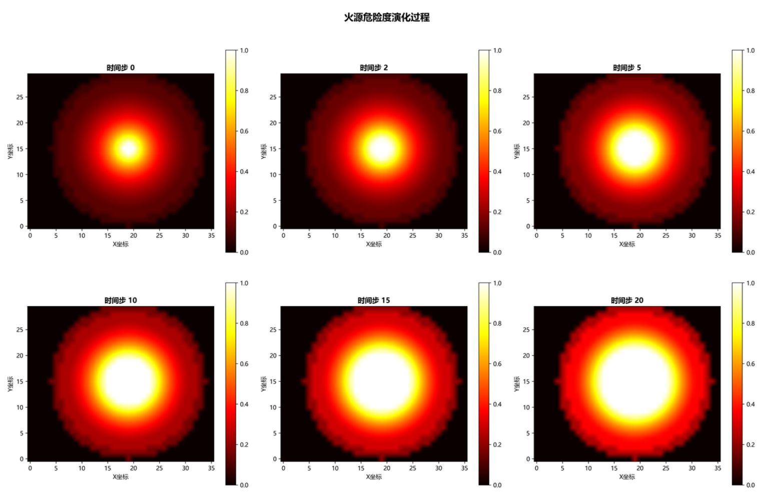

Evacuation Trajectories - Episode episode_1725
Exit
Robot Start
Robot Path
Robot End
Y Position
X Position
0
5
10
15
20
25
30
35

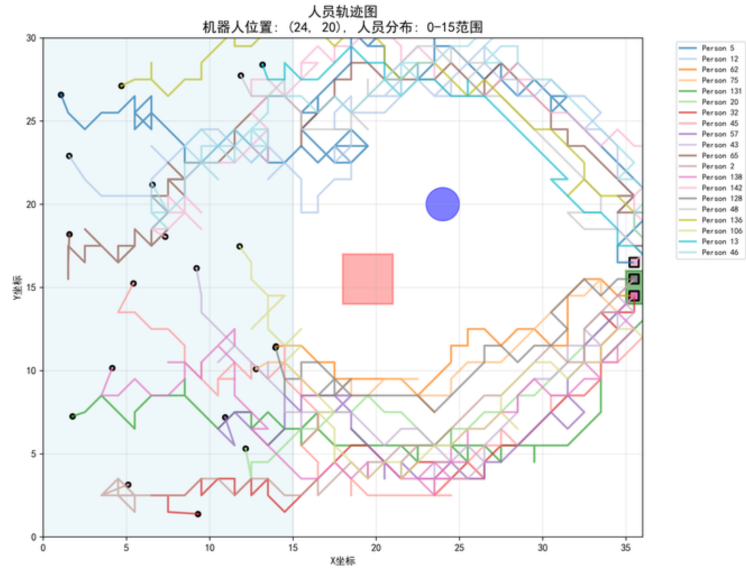
重构代码后发现问题所在：若人在关键位置死去就会变相提高疏散时间（健康值反向升高问题：假如出现了人死亡后剔除则会导致整体健康值上升不能表示目前平均状态）

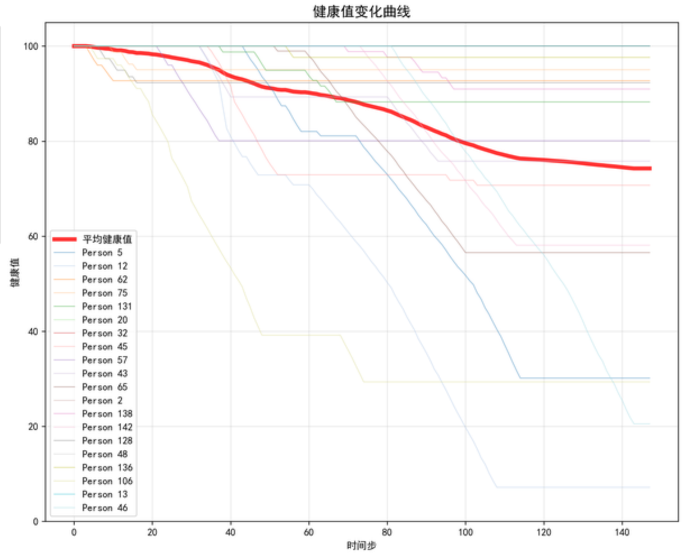
考虑拖行时：加入机器人会使得人的健康值上升但是时间同样上升的矛盾。

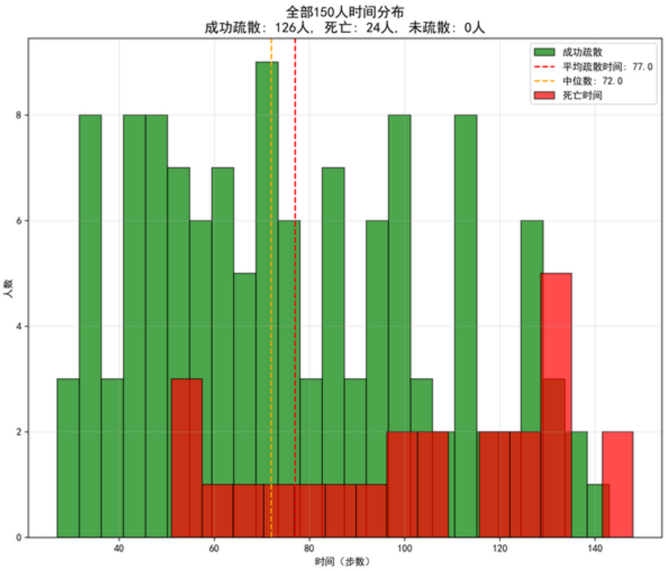


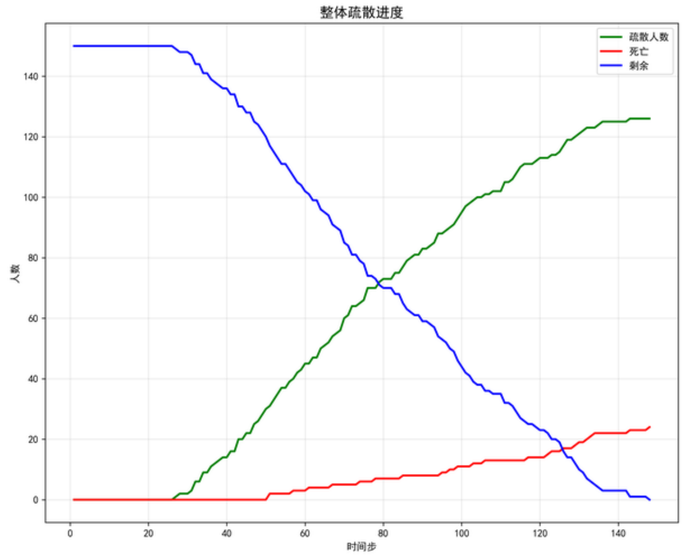



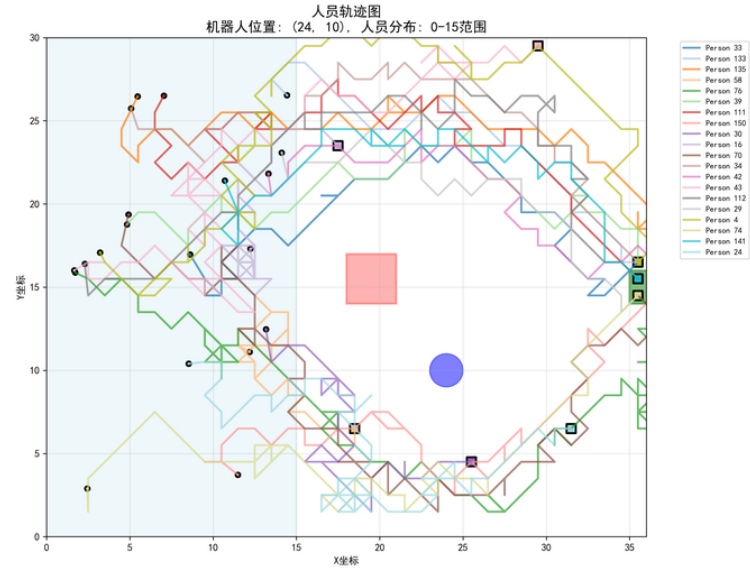



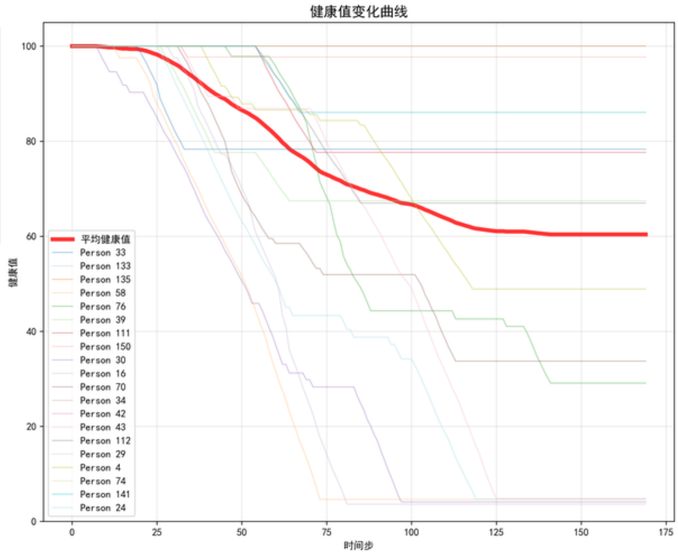

人员轨迹图
机器人位置：(24, 20)，人员分布：0-15范围
X坐标
Y坐标
Person 5
Person 12
Person 62
Person 75
Person 131
Person 20
Person 32
Person 45
Person 57
Person 43
Person 65
Person 2
Person 138
Person 142
Person 128
Person 48
Person 136
Person 106
Person 13
Person 46

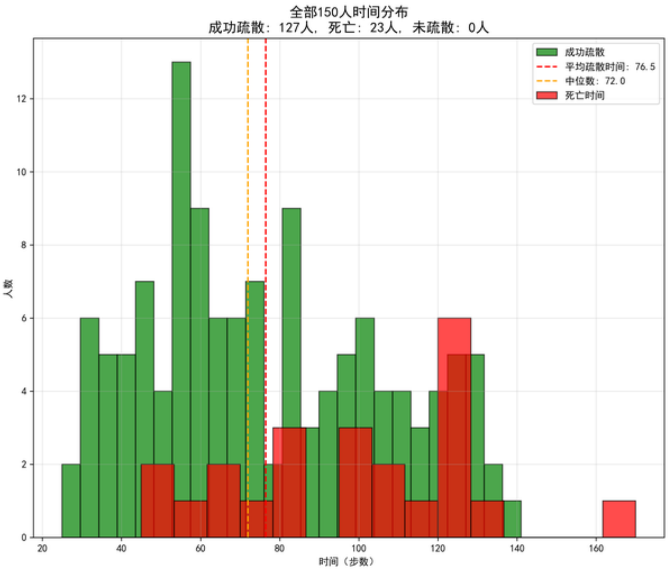

健康值变化曲线
健康值
时间步
平均健康值
Person 5
Person 12
Person 62
Person 75
Person 131
Person 20
Person 32
Person 45
Person 57
Person 43
Person 65
Person 2
Person 138
Person 142
Person 128
Person 48
Person 136
Person 106
Person 13
Person 46

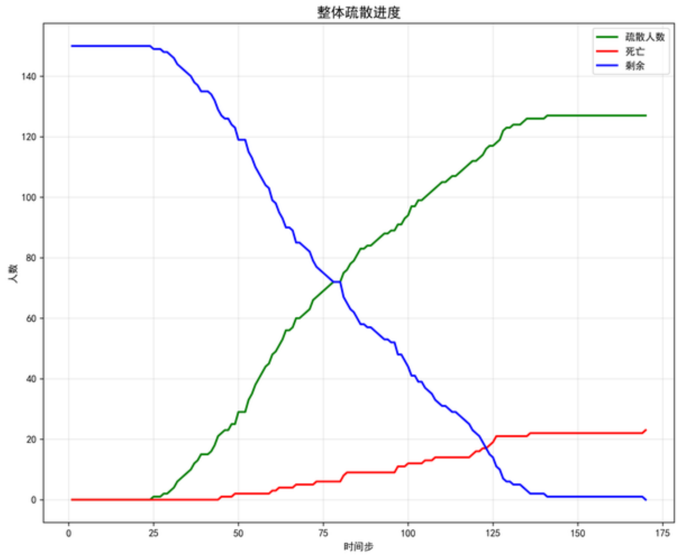

全部150人时间分布
成功疏散：126人，死亡：24人，未疏散：0人
成功疏散
平均疏散时间：77.0
中位数：72.0
死亡时间
人数
时间（步数）

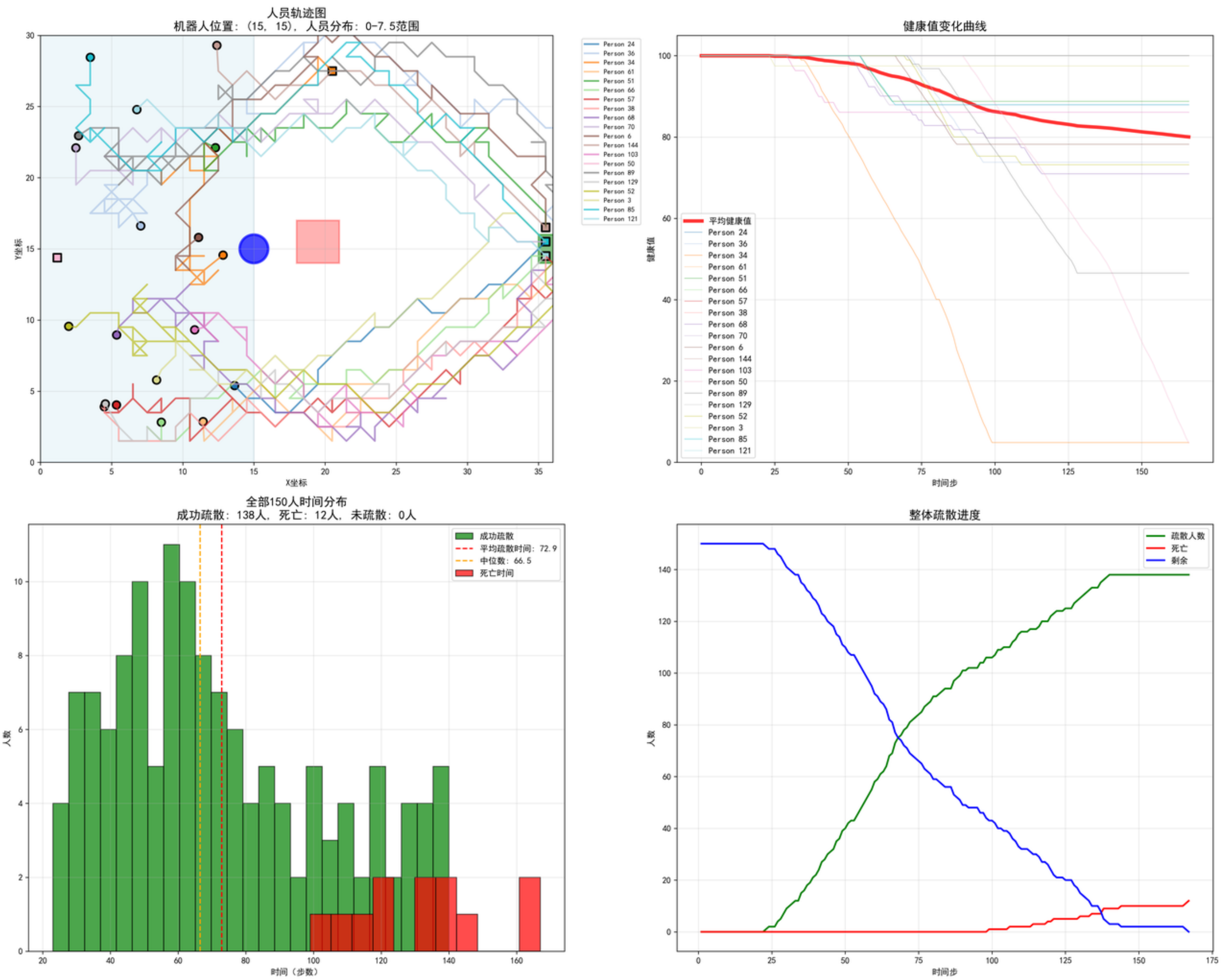

整体疏散进度
疏散人数
死亡
剩余
人数
时间步


人员轨迹图
机器人位置：(24, 10)，人员分布：0-15范围
Y坐标
X坐标
Person 33
Person 133
Person 135
Person 58
Person 76
Person 39
Person 111
Person 150
Person 30
Person 16
Person 70
Person 34
Person 42
Person 43
Person 112
Person 29
Person 4
Person 74
Person 141
Person 24

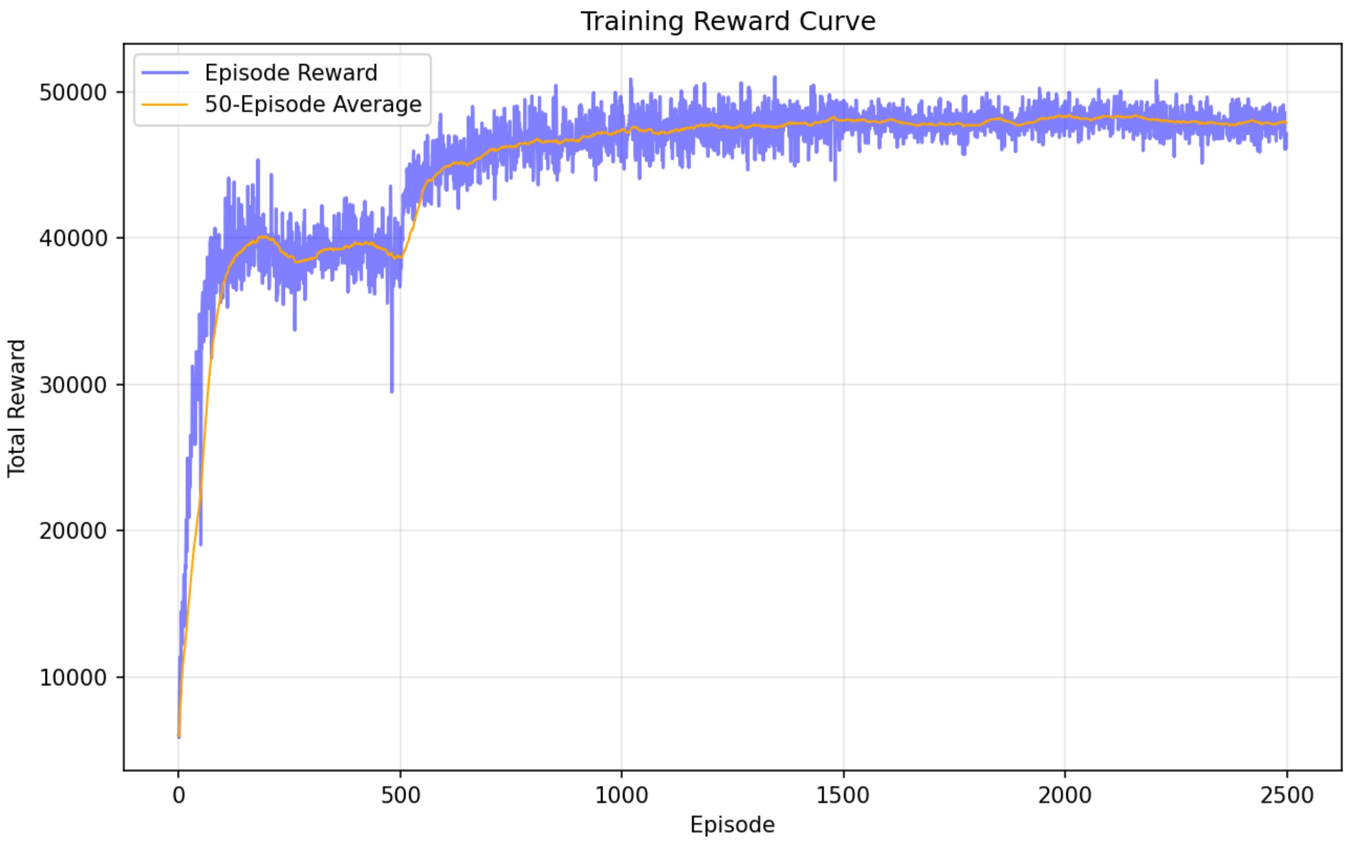

健康值变化曲线
健康值
时间步
平均健康值
Person 33
Person 133
Person 135
Person 58
Person 76
Person 39
Person 111
Person 150
Person 30
Person 16
Person 70
Person 34
Person 42
Person 43
Person 112
Person 29
Person 4
Person 74
Person 141
Person 24


全部150人时间分布
成功疏散：127人，死亡：23人，未疏散：0人
人数
时间（步数）
成功疏散
平均疏散时间：76.5
中位数：72.0
死亡时间

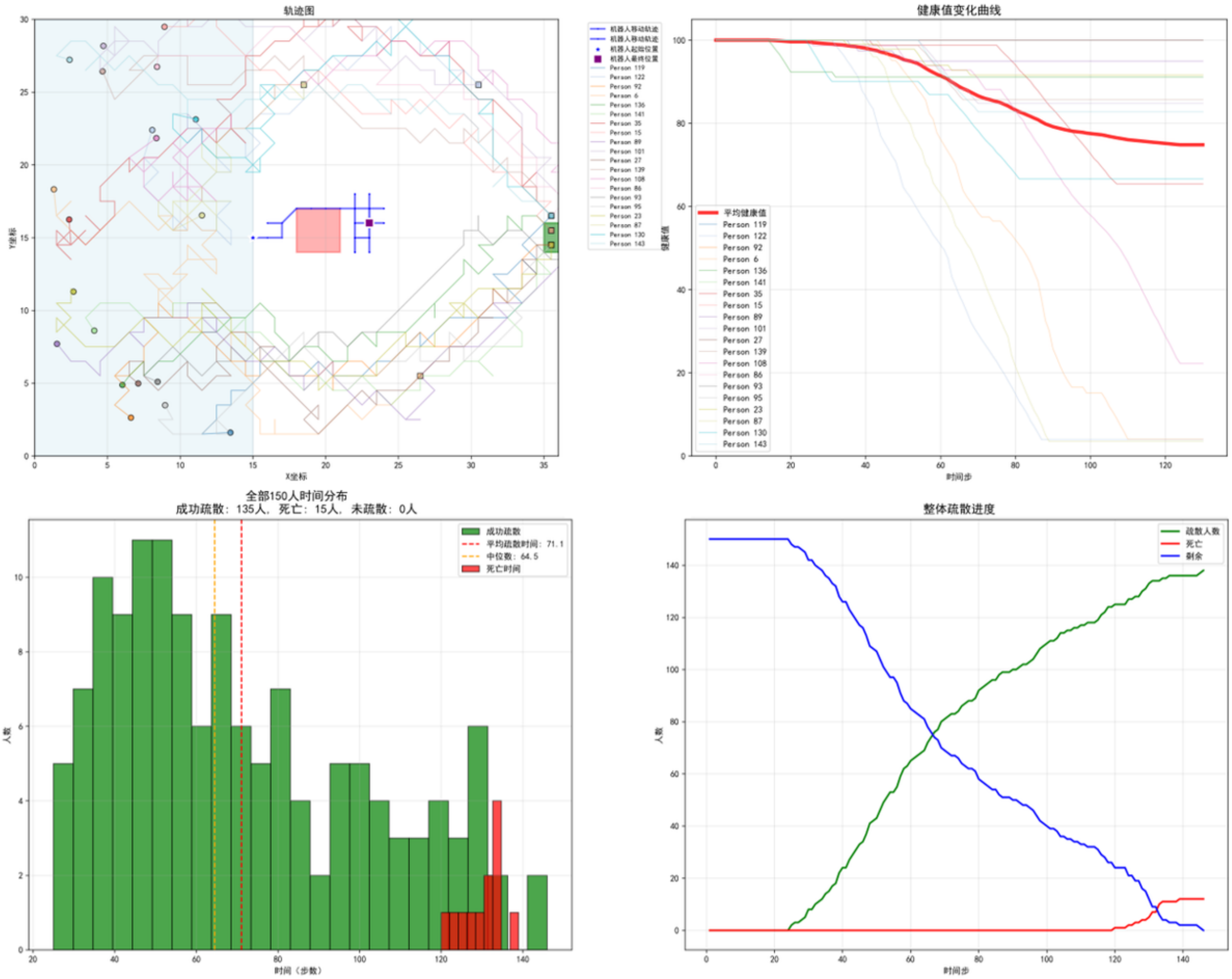

整体疏散进度
人数
时间步
疏散人数
死亡
剩余




| | | | | | | |
|---|---|---|---|---|---|---|
| 无机器人 | 139人 | 11人 | 92.7% | 7.3% | 146步 | 71.7 |
| 机器人(24,10) | 127人 | 23人 | 84.7% | 15.3% | 170步 | 72.4 |
| 机器人(24,20) | 126人 | 24人 | 84% | 16% | 148步 | 74.3 |
| 机器人（15,15） | 138 | 12人 | 92% | 8% | 162 步 | 78.2 |
| 机器人DRL | 143人 | 7 人 | 95.3% | 4.7% | 142步 | 64.4 |


Training Reward Curve
Episode Reward
50-Episode Average
Total Reward
Episode
50000
40000
30000
20000
10000
0
500
1000
1500
2000
2500

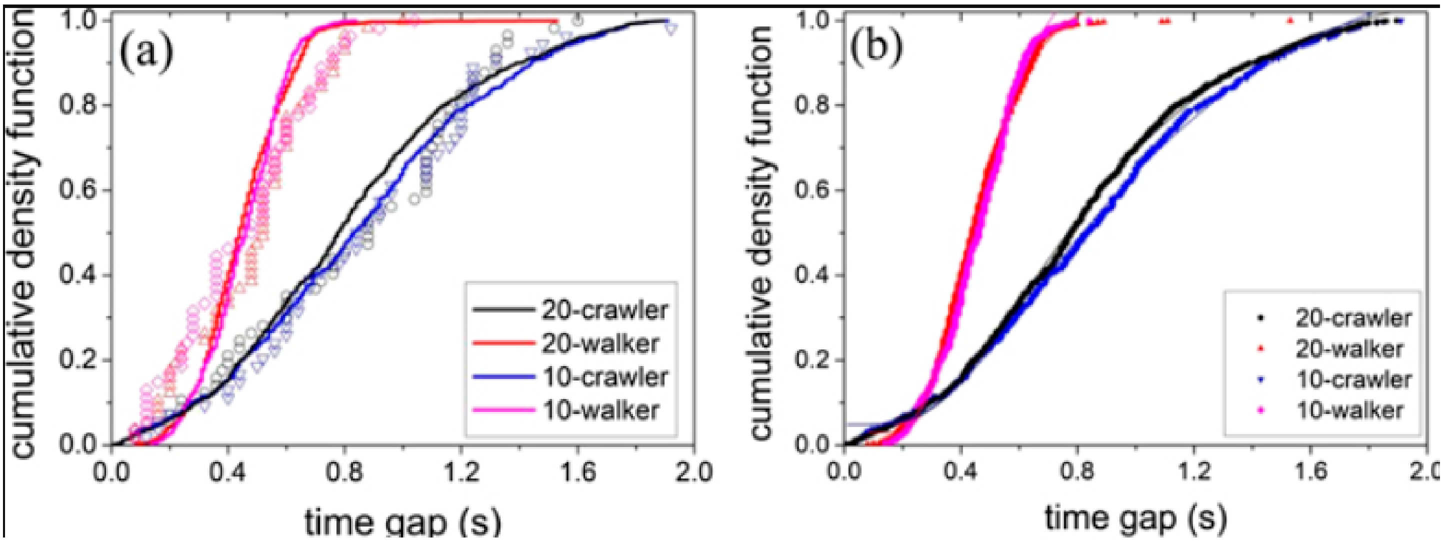
经过训练后




爬行：多用实验情况得到结果

1.合肥工业大学20人小实验（结果已简化）实验环境类似于我们的环境

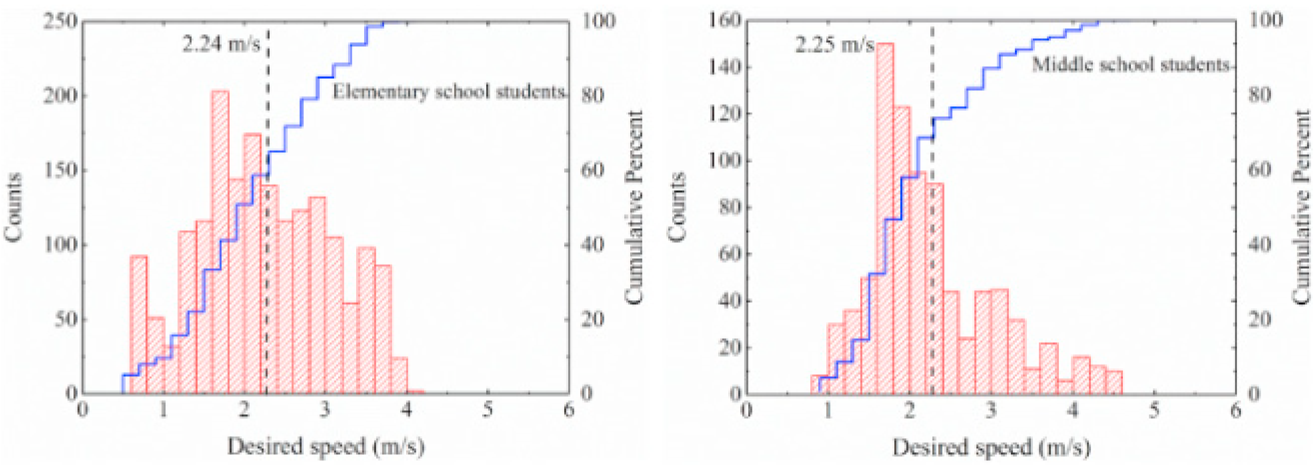
https://iopscience.iop.org/article/10.1088/1742-5468/abe945




2.中小学学生速度 if4.5

https://www.sciencedirect.com/science/article/pii/S2212420919317327

First of all, we measure the average desired speed of all students in individual motion. Fig. 4 shows the relative frequency and cumulative frequency distributions of the desired speed for different age group students in UW area. The average UW desired speeds for elementary and middle school students are 2.24 and 2.25 m/s, and there is no significant statistical difference between them by means of T-test ($t=-0.28$, $p=0.78$). The same results are obtained from the distributions of the KHC desired speed as illustrated in Fig. 5. The average KHC desired speeds are 1.02 and 1.03 m/s for elementary and middle school students, respectively. Compared with the average desired speed of UW, the average KHC desired speed reduces by 54%.



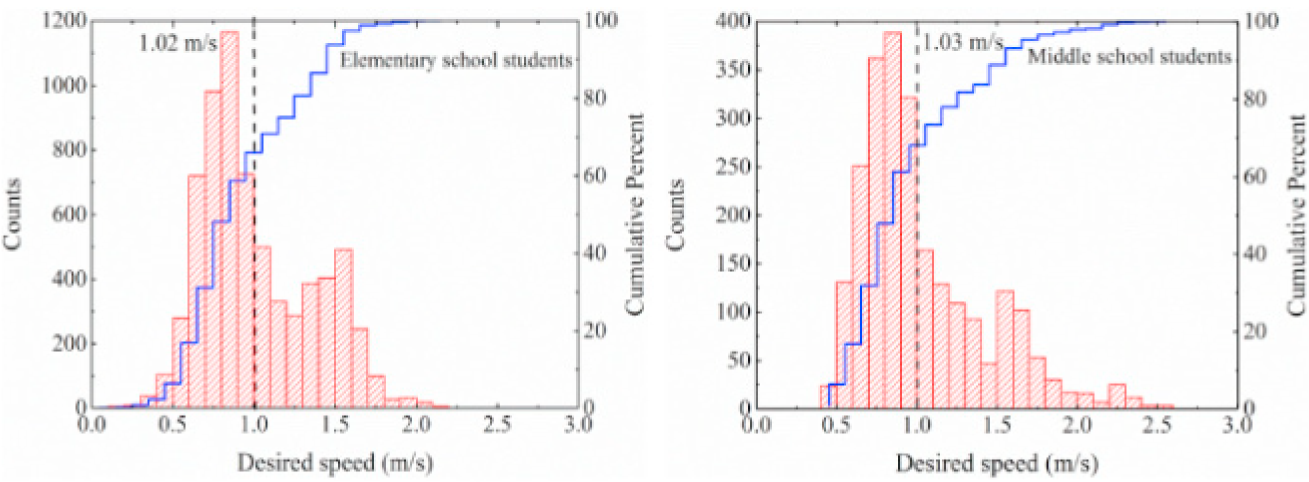

(a) Elementary school students

(b) Middle school students

Download: Download high-res image (1MB)

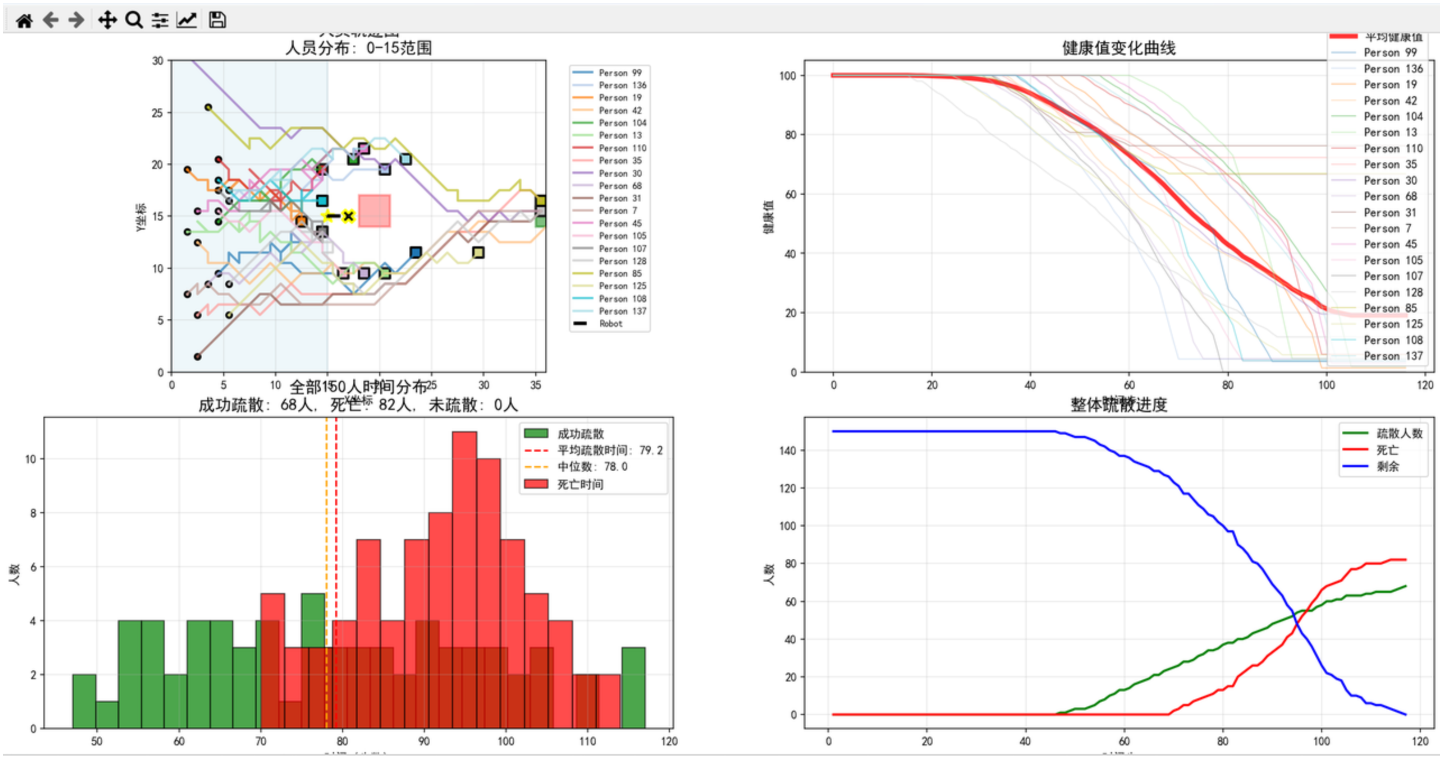
Download: Download full-size image

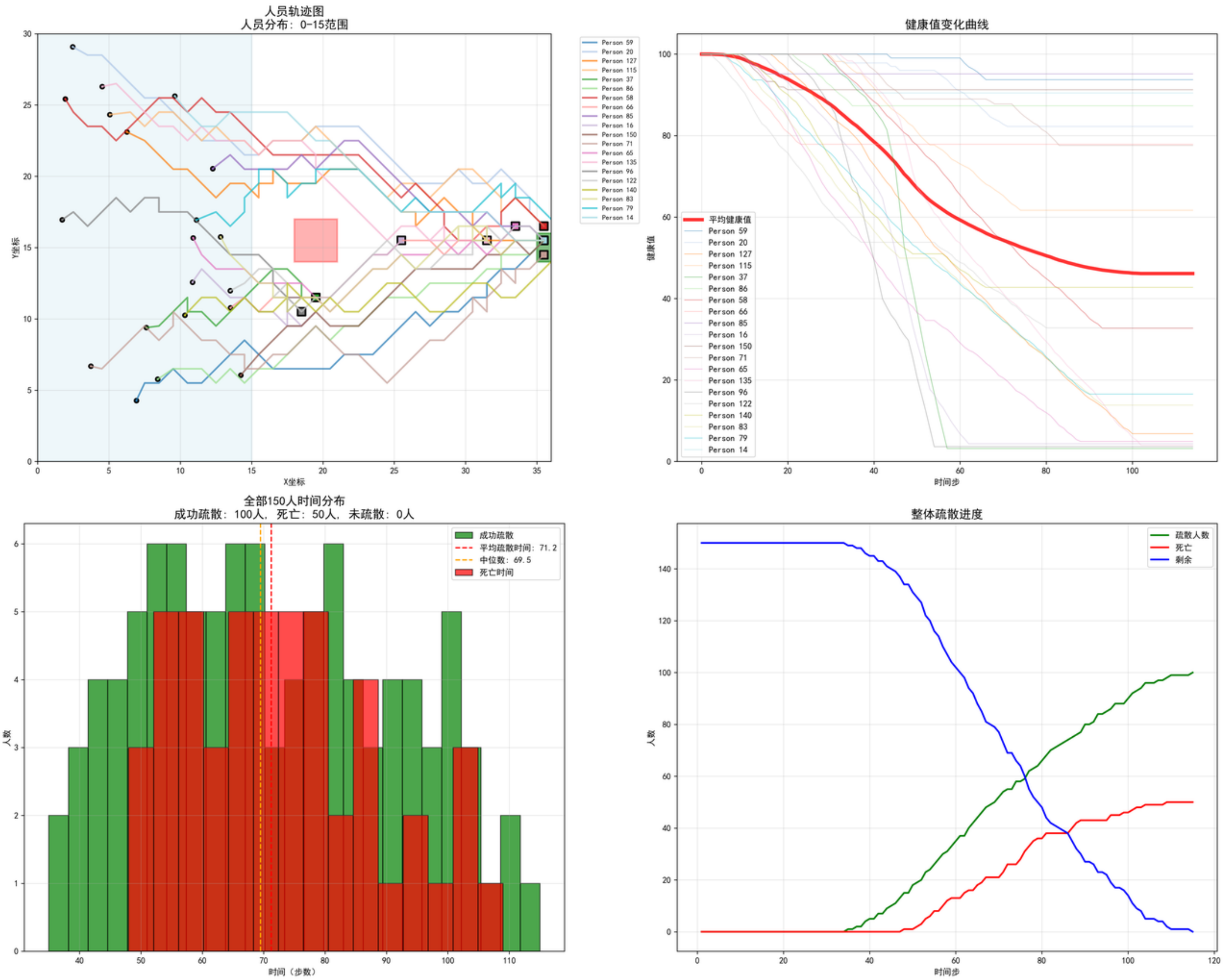
Fig. 4. Distribution of the desired speed in UW area. The average UW desired speeds for elementary and middle school students are 2.24 and 2.25 m/s, respectively.




(a) Elementary school students

(b) Middle school students


人员分布：0-15范围
Y坐标
X坐标
Robot
健康值变化曲线
平均健康值
健康值
全部150人时间分布
成功疏散：68人，死亡：82人，未疏散：0人
成功疏散
平均疏散时间：79.2
中位数：78.0
死亡时间
人数
整体疏散进度
疏散人数
死亡
剩余

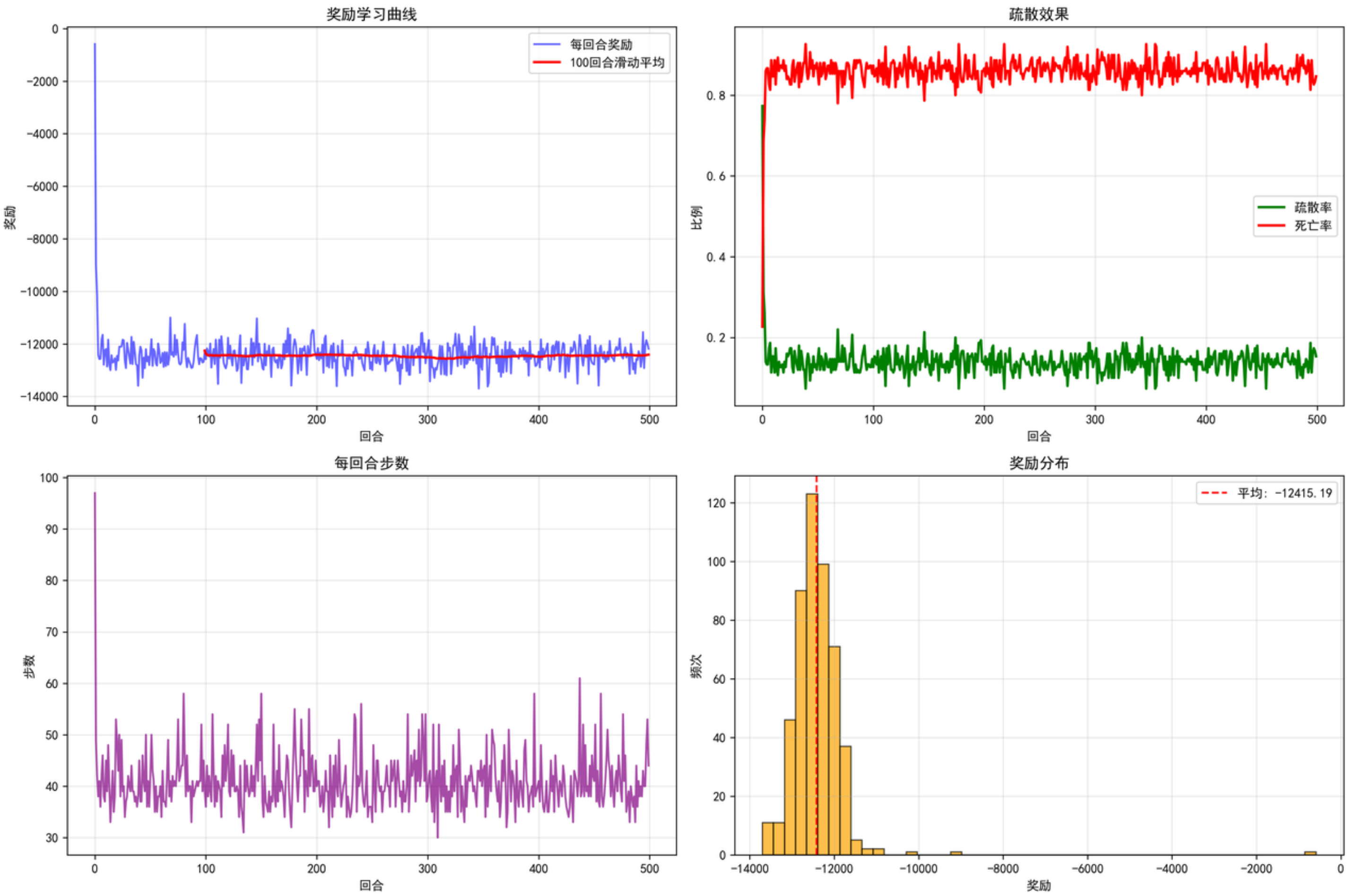

人员轨迹图
人员分布：0-15范围
Y坐标
X坐标
健康值变化曲线
平均健康值
健康值
时间步
全部150人时间分布
成功疏散：100人，死亡：50人，未疏散：0人
成功疏散
平均疏散时间：71.2
中位数：69.5
死亡时间
人数
时间（步数）
整体疏散进度
疏散人数
死亡
剩余
时间步

```python
class Person:

    # 速度常量
    NORMAL_SPEED = 1.25    # m/s，整体行进速度降低
    CRAWLING_SPEED = 0.6   # m/s，当健康值极低时的爬行速度

    def __init__(self, id, pos_x, pos_y):
        self.id = id
        self.pos = (pos_x, pos_y)
        self.speed = Person.NORMAL_SPEED
```

roblems    Output    Debug Console    Terminal    Ports    powershe

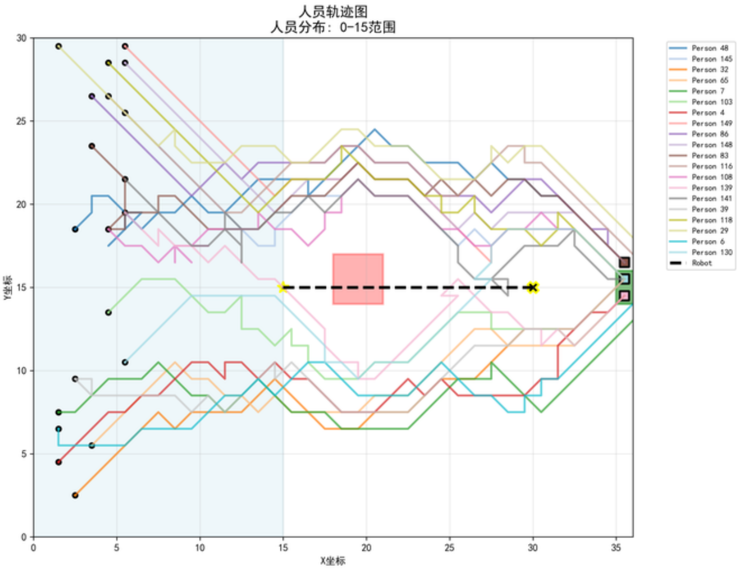
```
:\software\miniconda\envs\AC\lib\site-packages\numpy\core\_methods.py:192: RuntimeWarning: invalid
calar divide
 ret = ret.dtype.type(ret / rcount)
pisode   50: Reward=-14809.26, Avg100=-14534.70, Steps= 56, Evac=0.00%, Death=100.00%, ε=0.2592
pisode   99: Reward=-14825.62, Avg100=-14580.33, Steps= 52, Evac=0.00%, Death=100.00%, ε=0.0708
```

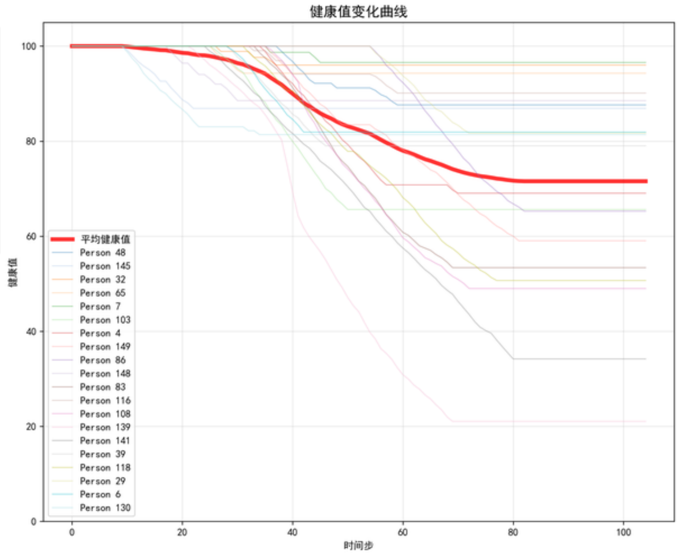
爬行速度牵一发而动全身，稍微降一点死亡率就变成100%，并且爬行会挡路，健康值越少速度越慢，并且人不容易出去，出口就一个，一个爬行很容易堵在出口，并且通过结果调试爬行健康值也会极端，本身速度已经降低了，放一个出来就会像多米洛骨牌一样都爬，死亡人数会大增。

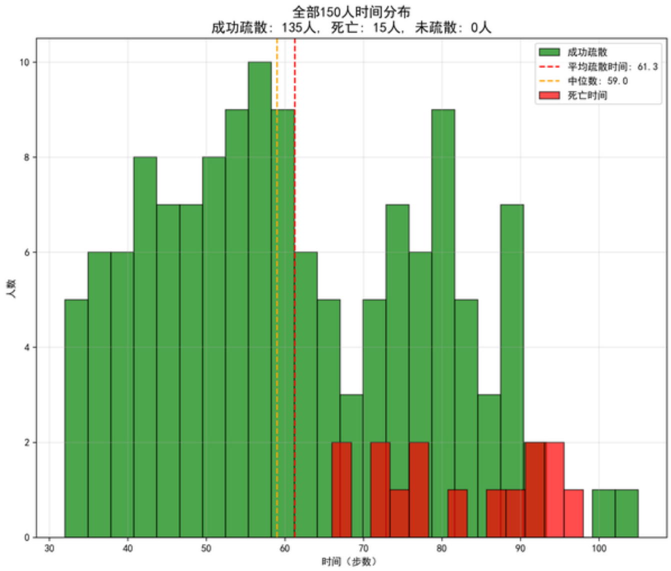


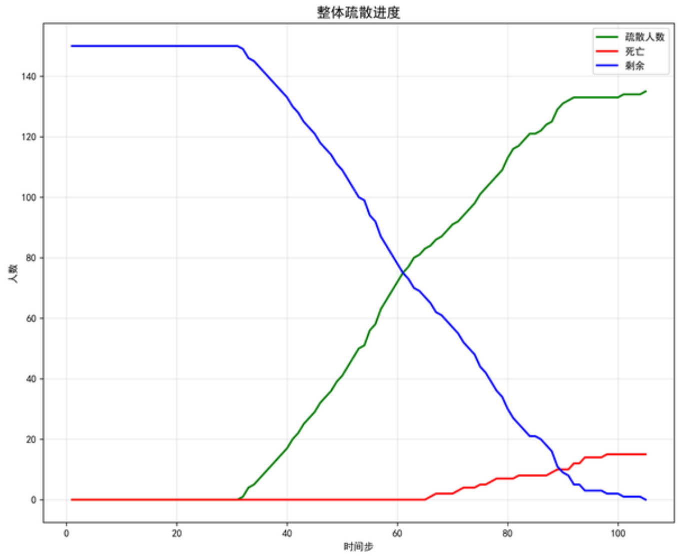

把行人更细致地分布后效果变好一点

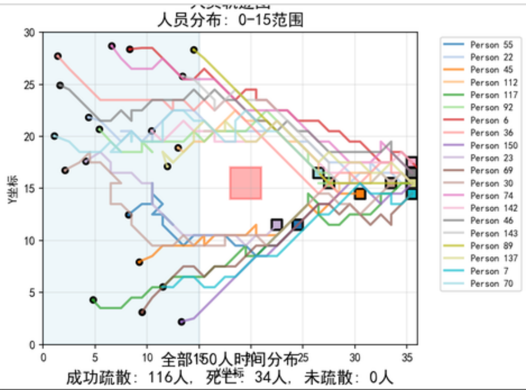
1.机器人离火太近了，没有移动空间，原文中的火离人员会更远。2.机器人不能从火上趟过去

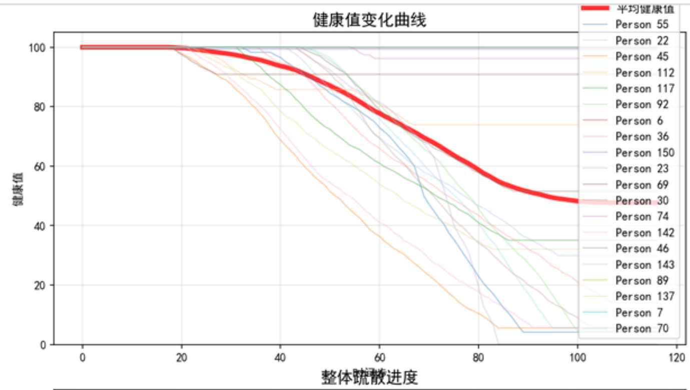

人员轨迹图
人员分布：0-15范围
X坐标
Y坐标
Person 48
Person 145
Person 32
Person 65
Person 7
Person 103
Person 4
Person 149
Person 86
Person 148
Person 83
Person 116
Person 108
Person 139
Person 141
Person 39
Person 118
Person 29
Person 6
Person 130
Robot

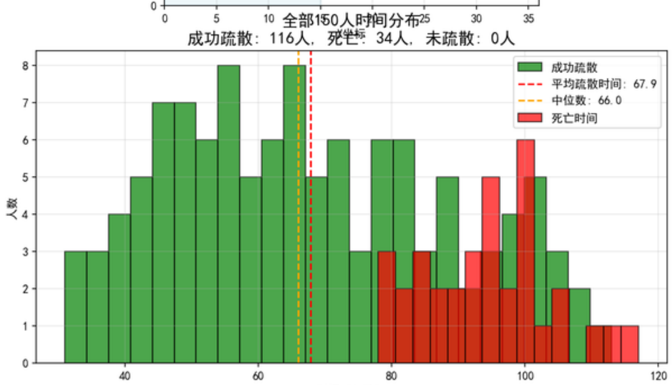

健康值变化曲线
健康值
时间步
平均健康值
Person 48
Person 145
Person 32
Person 65
Person 7
Person 103
Person 4
Person 149
Person 86
Person 148
Person 83
Person 116
Person 108
Person 139
Person 141
Person 39
Person 118
Person 29
Person 6
Person 130

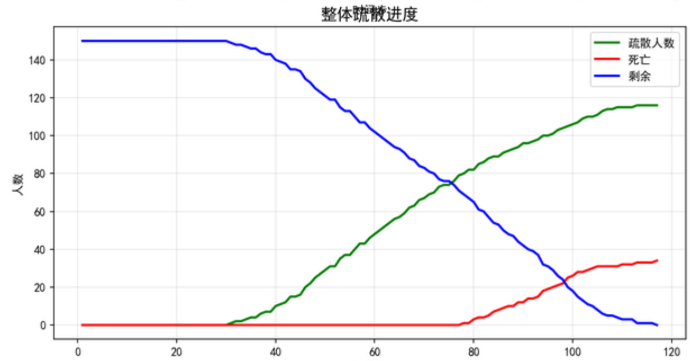

全部150人时间分布
成功疏散：135人，死亡：15人，未疏散：0人
成功疏散
平均疏散时间：61.3
中位数：59.0
死亡时间
人数
时间（步数）


整体疏散进度
疏散人数
死亡
剩余
人数
时间步


人员分布：0-15范围
Y坐标
Person 55
Person 22
Person 45
Person 112
Person 117
Person 92
Person 6
Person 36
Person 150
Person 23
Person 69
Person 30
Person 74
Person 142
Person 46
Person 143
Person 89
Person 137
Person 7
Person 70


健康值变化曲线
健康值
Person 55
Person 22
Person 45
Person 112
Person 117
Person 92
Person 6
Person 36
Person 150
Person 23
Person 69
Person 30
Person 74
Person 142
Person 46
Person 143
Person 89
Person 137
Person 7
Person 70


成功疏散：116人，死亡：34人，未疏散：0人
成功疏散
平均疏散时间：67.9
中位数：66.0
死亡时间
人数

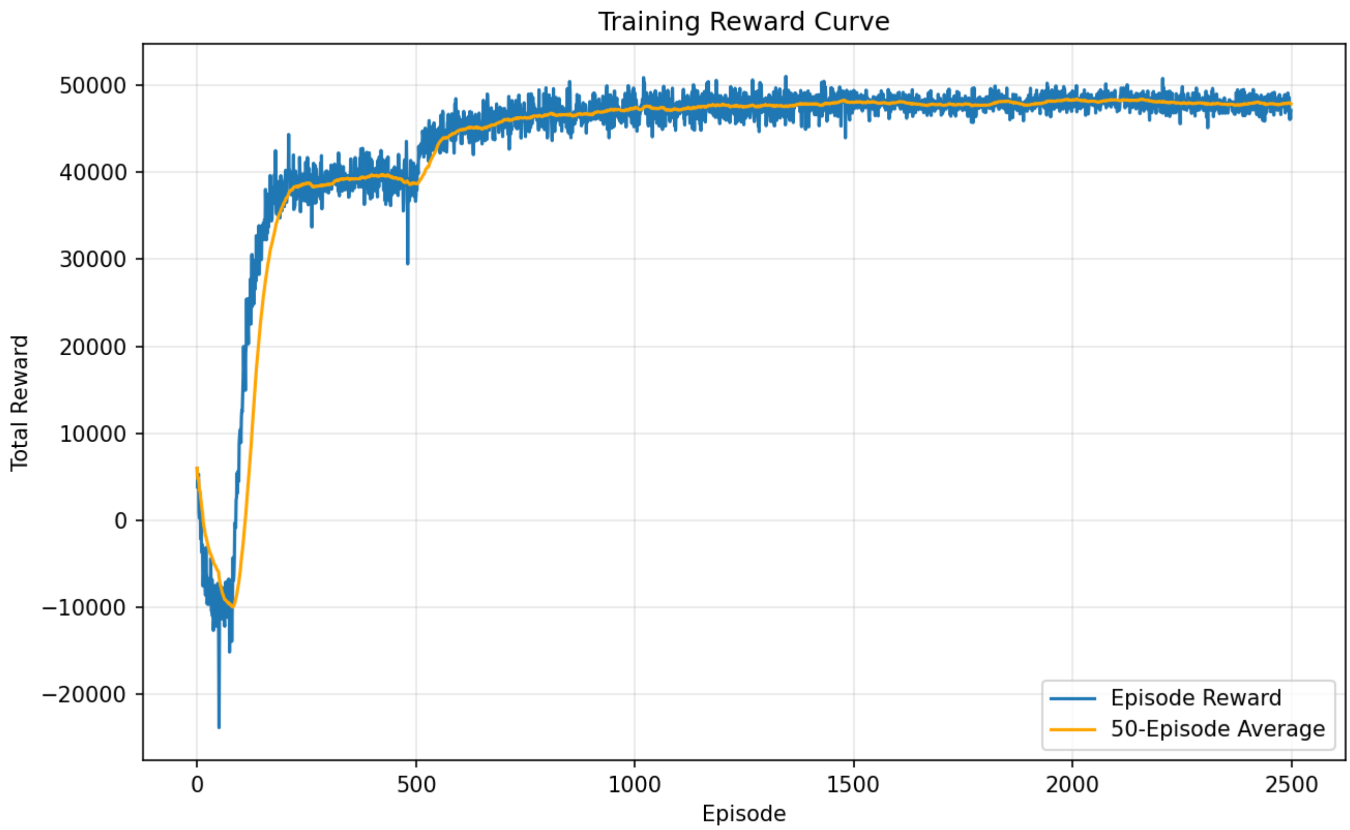

整体疏散进度
疏散人数
死亡
剩余
人数

| 策略 | 平均健康值 | 平均疏散时间 |
|---|---|---|
| 无机器人 | 78.34 | 412.7 |
| 静态机器人 | 71.12 | 458.2 |
| DQN 机器人 | 86.95 | 395.4 |

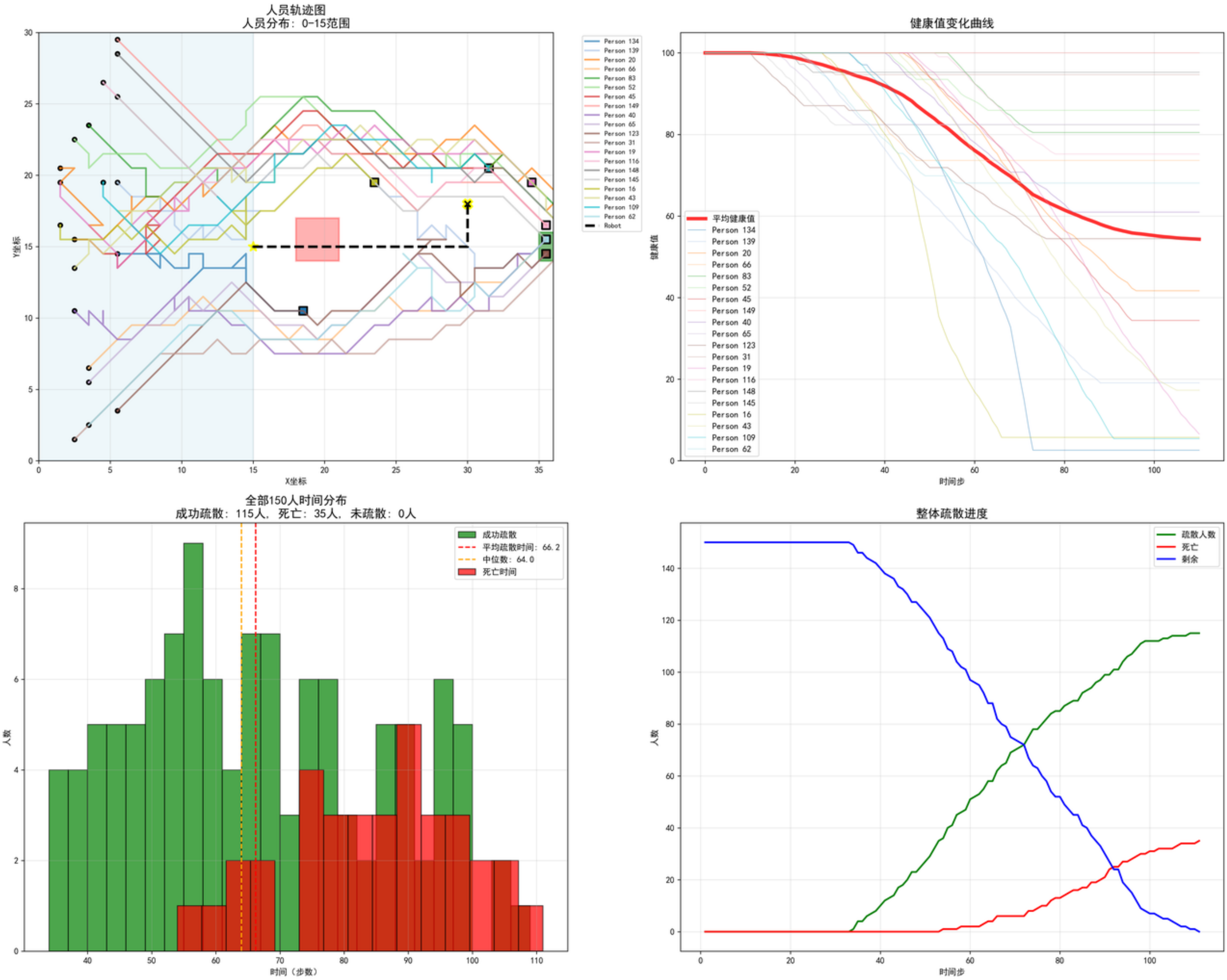
改变了行人的行为逻辑，让健康值最低的人先走而不是随机行走。

dqn：开始在火源前移动是的人尽可能远离火，速度更快，后期在火源后面来回疏散避免人扎堆，通过减慢爬行速度从而放置因人员在出口拥挤导致的死亡。

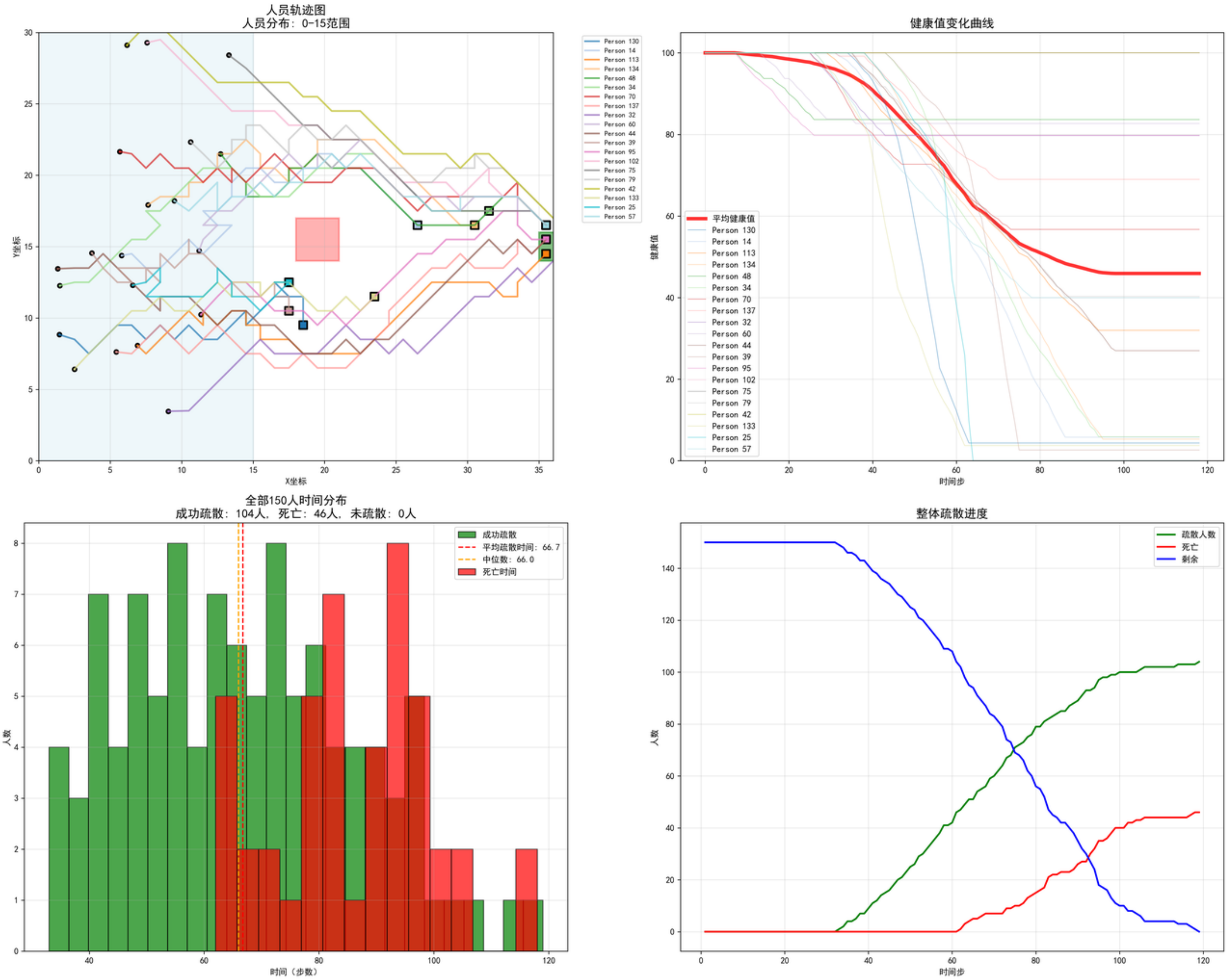
dqn机器人 平均健康值：52.6 时间53.9s



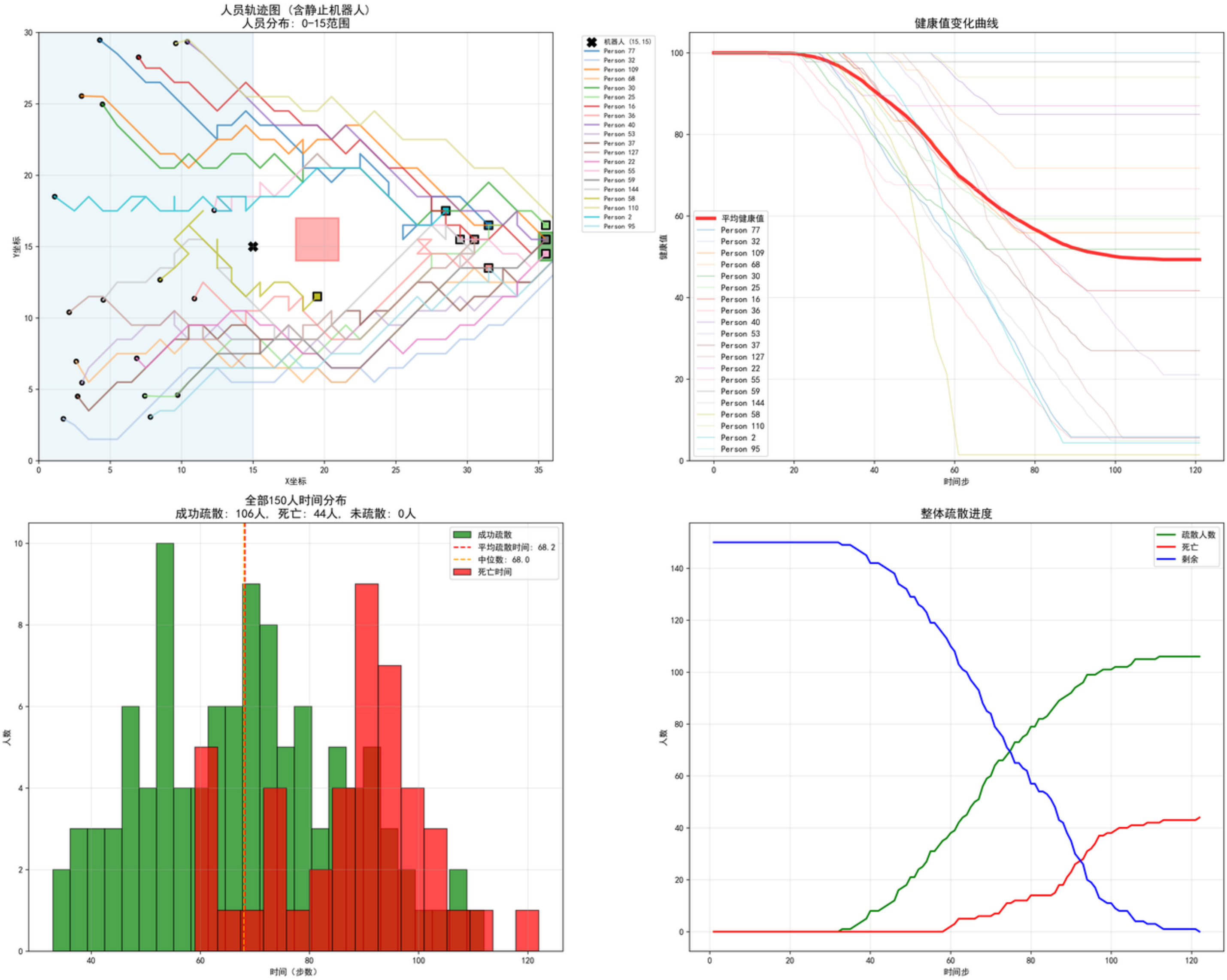

dqn机器人:时间：55.5s 最终健康值: 平均=54.2




dqn机器人:平均健康值=54.6 时间：53.5s




静态机器人：平均健康值：48.9 时间：59.2s




无机器人：平均健康值：46.7 时间：58.9s

主要查的论文也是关于健康值与速度之间的关系，尤其是健康值很低的时候与速度之间的对应关系

特殊场景：医院，封闭商场

温度>50°C时，速度衰减率与温度呈指数关系：
vT=v0×e－0.021(T－50)

能见度>10m时：速度接近正常值

能见度5-10m时：v = 1.22 - 0.072V（每降1m速度衰减7.2%）

能见度<5m时：速度骤降至正常值的35%，且跌倒概率增加300%

临界阈值：能见度≤3m时，决策时间延长2.1倍，疏散效率下降68%

D.A. Purser, 2002. Toxicity Assessment of Combustion Products, SFPE Handbook of Fire Protection Engineering, National Fire Protection Association, Quincy, U. K.量化危险产物

加入**1**、决策时间（？）

医院病房测试显示：温度60°C时，老年患者（移动能力III级）疏散时间增加58%

荧光地标

在火灾动态发展过程中，关于人员逃生的安全要求如下：（1）最小清晰高度以下的烟层/空气层温度不超过60℃[6]；（2）最小清晰高度以下的烟层/空气层能见度不小于10m[7]；（3）最小清晰高度以下的烟层/空气层一氧化碳浓度不超过500 ppm[8]。


[6] Florencio Fernánez-Alaiz,Ana Maria Castañón,Fernando Gómez-Fernández,Antonio Bernardo-Sánchez,Marc Bascompta. Analysis of the Fire Propagation in a Sublevel Coal Mine[J]. Energies,2020,13(14):

[7] NFPA 92B, 1991. Guide for Smoke Management Systems in Malls, Atria and Large Areas, National Fire Protection Association, Quincy, U. K.

[8] D.A. Purser, 2002. Toxicity Assessment of Combustion Products, SFPE Handbook of Fire Protection Engineering, National Fire Protection Association,Quincy, U. K.


会存在相似的事情：失能无法移动也无法干涉环境

**状态2**.“健康”、“受伤”、“重伤”、“濒死（失能）”受伤后移动速度减缓，重伤移动缓慢，濒死只能原地等待救援

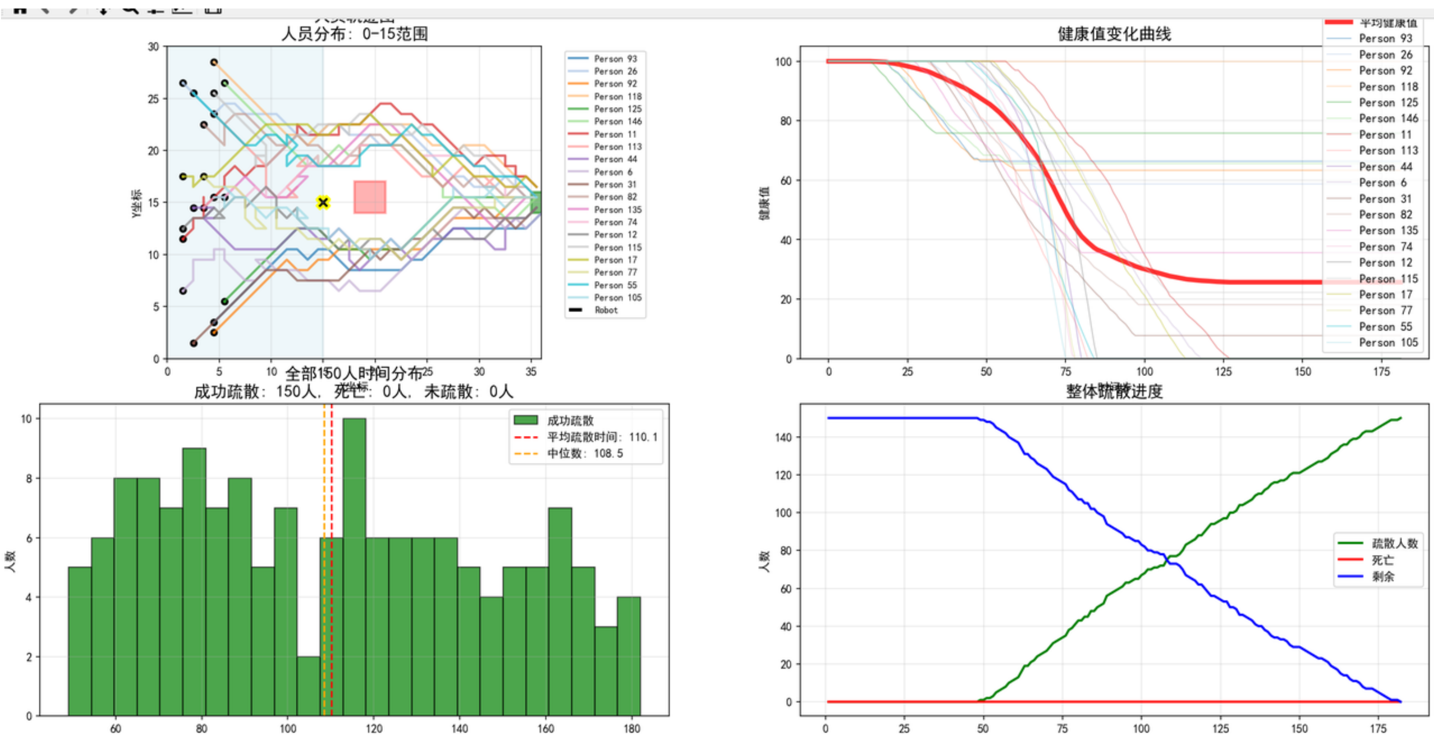
变化：1.将火灾与健康值之间的调控用非线性方式重构，取消死亡这一概念，优化健康值对速度的函数。

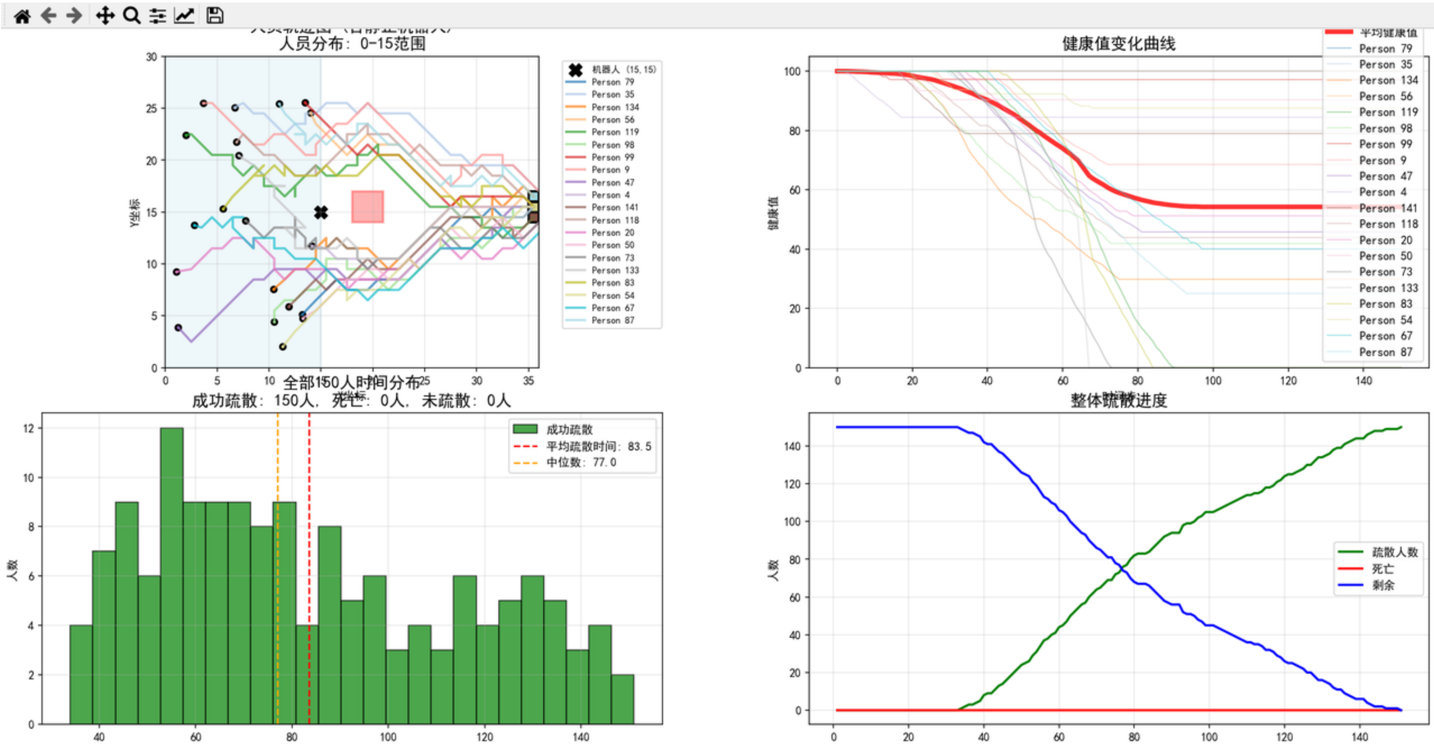
2.1奖励函数方面：个体奖励：加重关于疏散时间对奖励影响的权重（原奖励函数中疏散奖励和死亡惩罚应取消，这样奖励函数变得单一）[此刻想法是疏散80%人数时间（原方法）或者规定时间疏散人数]

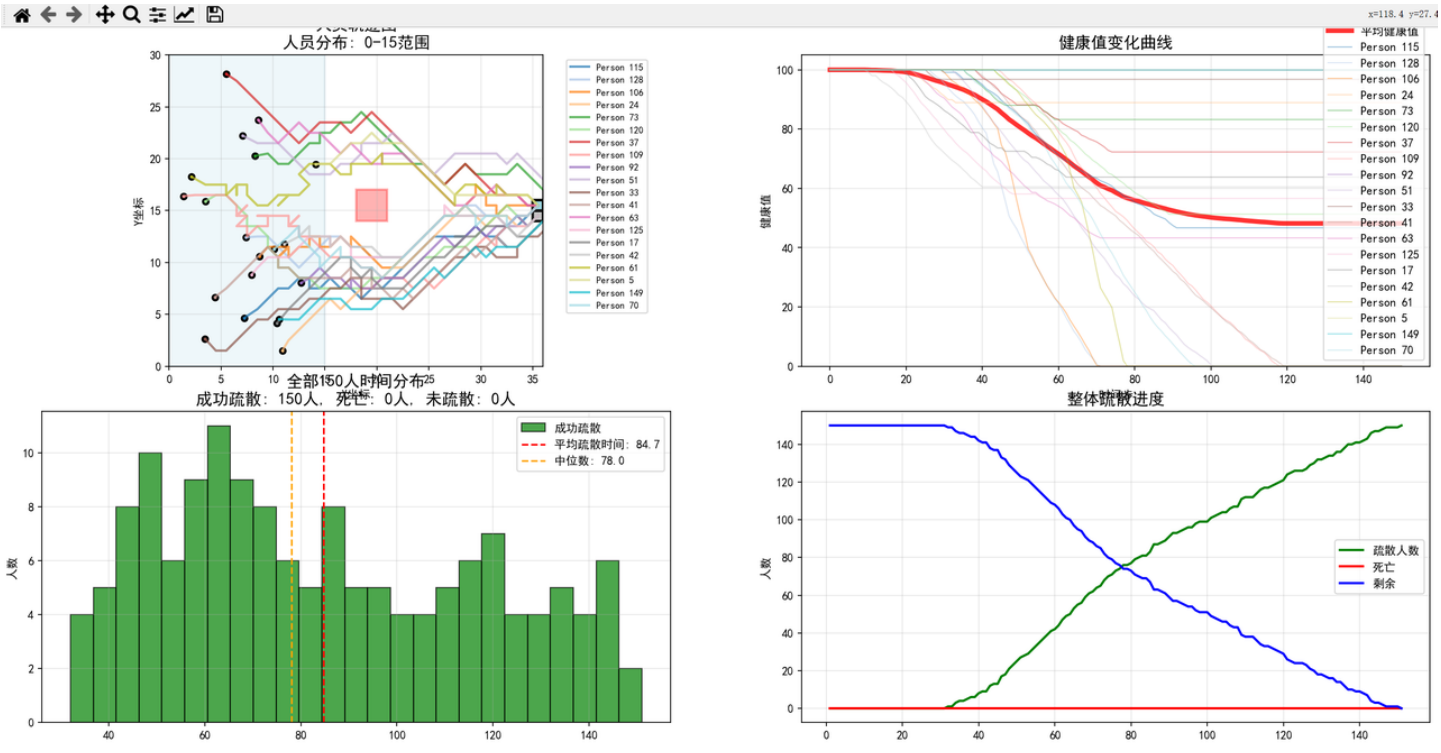
2.2群体奖励加入的东西暂时定为时间和协作奖励（避免两个机器人重复引导同一人群）

现在简单训练后发现加入dqn之后与无机器人效果近似

假如没有死亡：

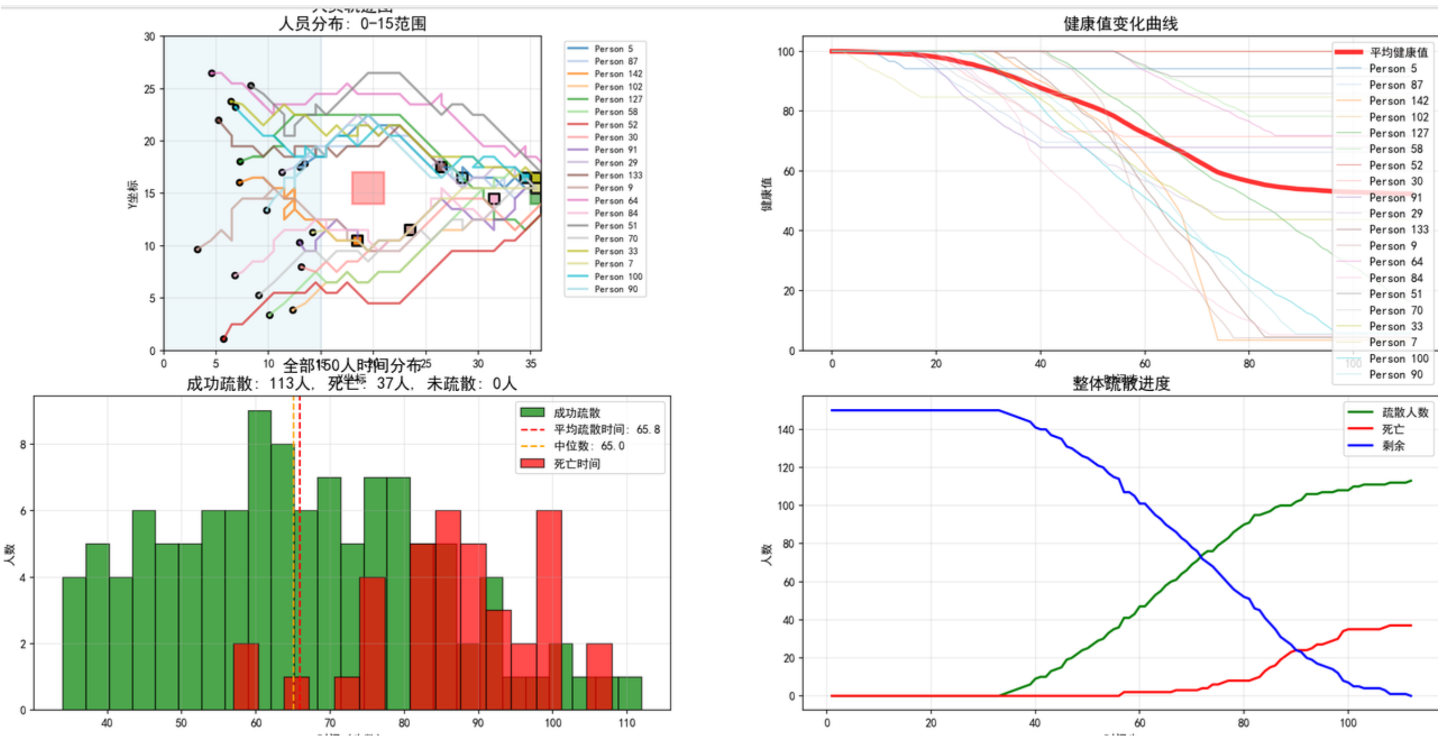

人员分布：0-15范围
Y坐标
Person 93
Person 26
Person 92
Person 118
Person 125
Person 146
Person 11
Person 113
Person 44
Person 6
Person 31
Person 82
Person 135
Person 74
Person 12
Person 115
Person 17
Person 77
Person 55
Person 105
Robot
健康值变化曲线
健康值
平均健康值
全部150人时间分布
成功疏散：150人，死亡：0人，未疏散：0人
人数
成功疏散
平均疏散时间：110.1
中位数：108.5
整体疏散进度
人数
疏散人数
死亡
剩余


人员分布：0-15范围
Y坐标
机器人 (15, 15)
Person 79
Person 35
Person 134
Person 56
Person 119
Person 98
Person 99
Person 9
Person 47
Person 4
Person 141
Person 118
Person 20
Person 50
Person 73
Person 133
Person 83
Person 54
Person 67
Person 87
健康值变化曲线
健康值
平均健康值
全部150人时间分布
成功疏散：150人，死亡：0人，未疏散：0人
人数
成功疏散
平均疏散时间：83.5
中位数：77.0
整体疏散进度
人数
疏散人数
死亡
剩余

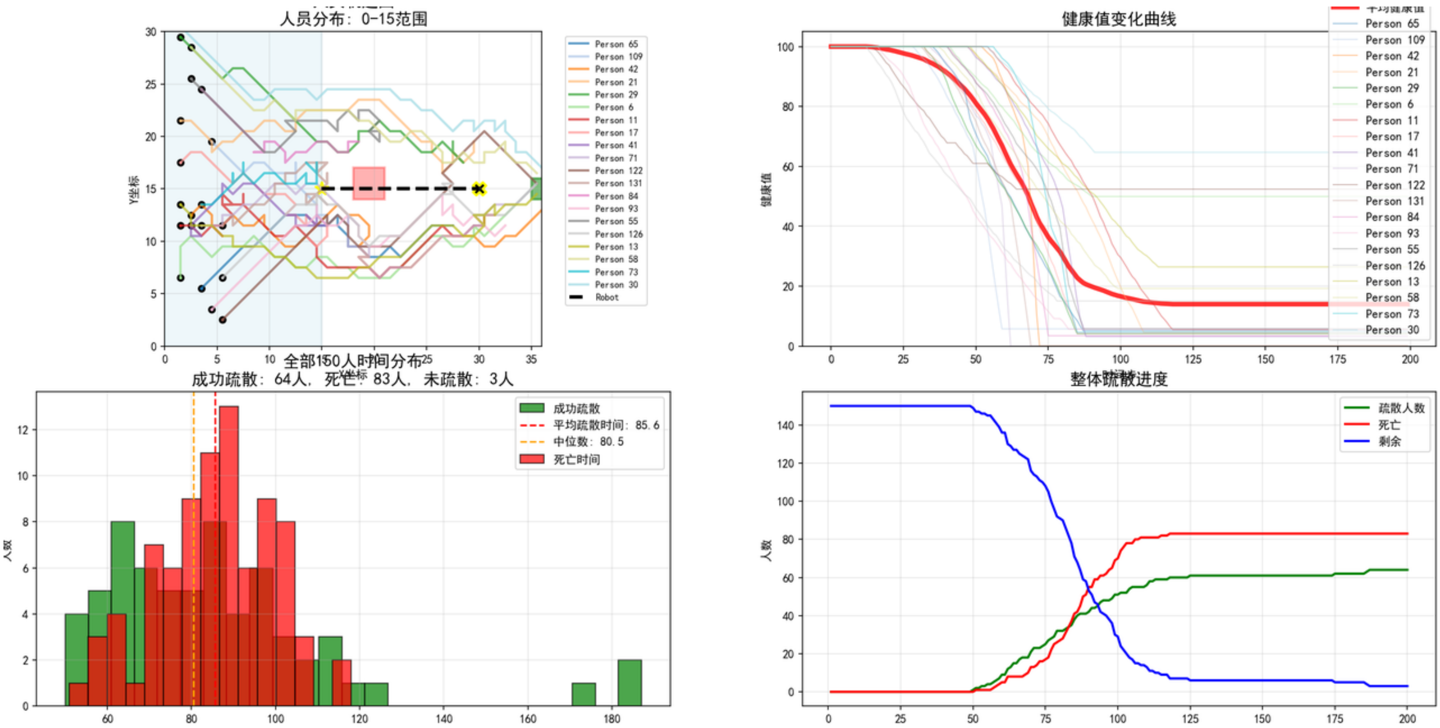



考虑从旁边绕过：

无机器人：

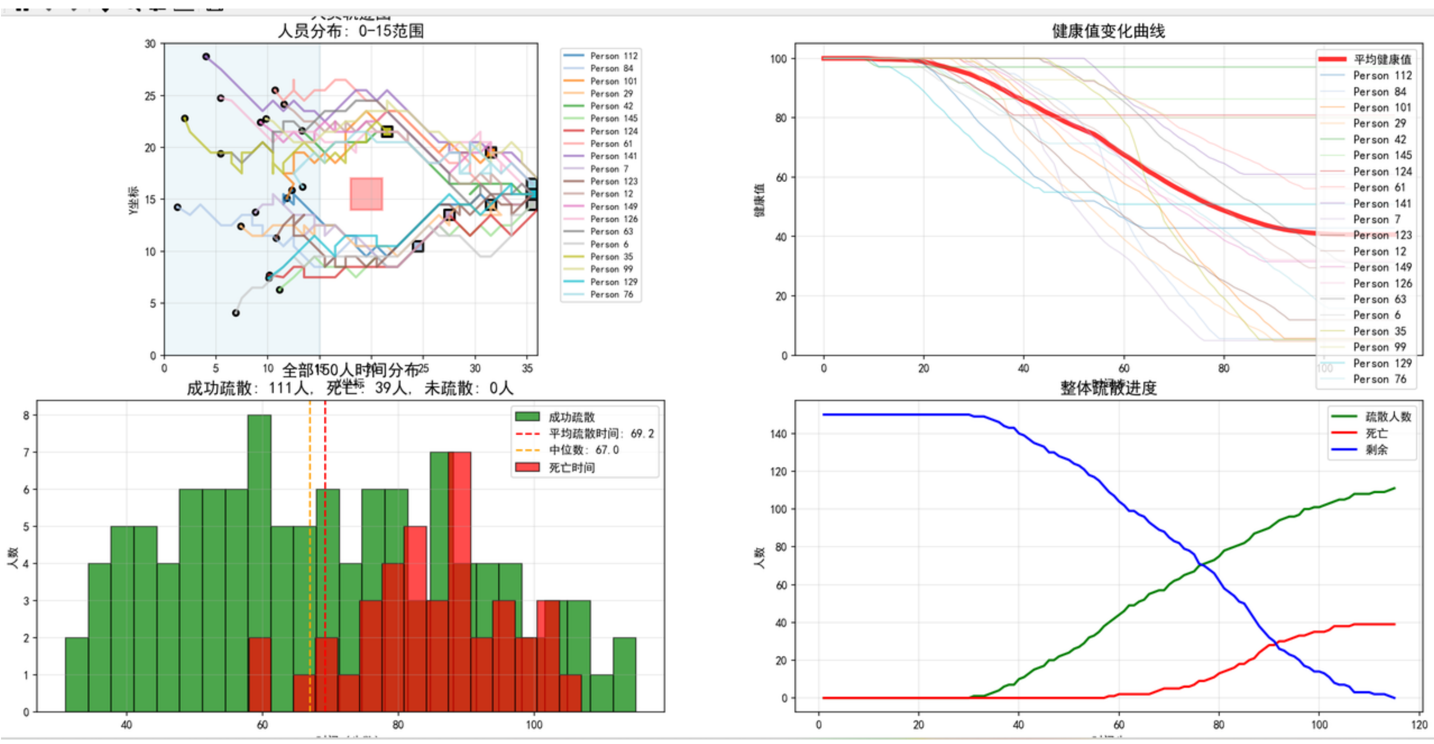



有dqn：可能是机器人与死亡人员全部堆在出口更影响出去效率

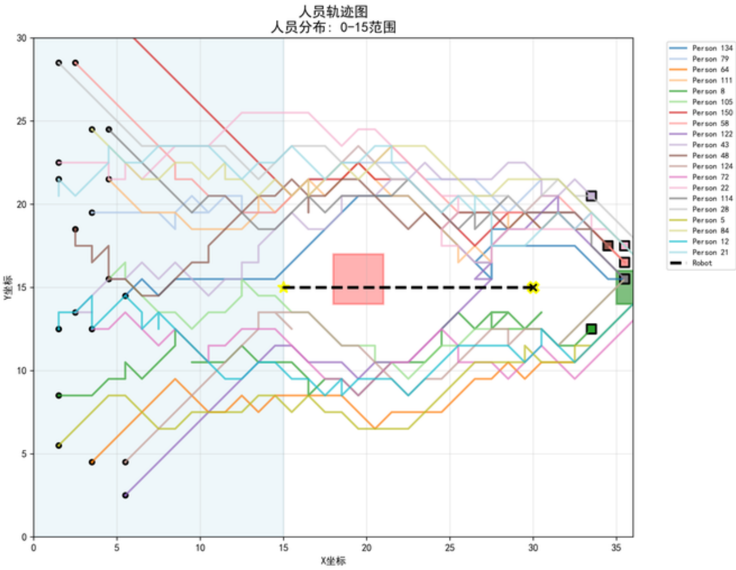


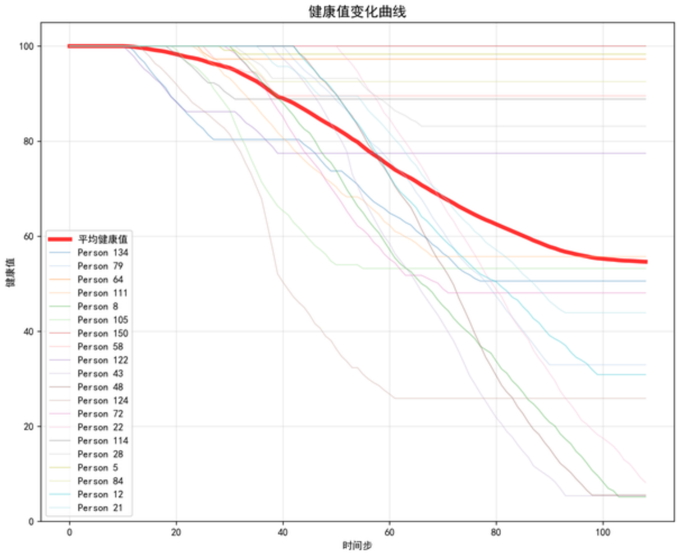

死亡后踩踏降低速度：降低火灾风险度之后死亡率也明显下降（怀疑是最后一步影响不大）、

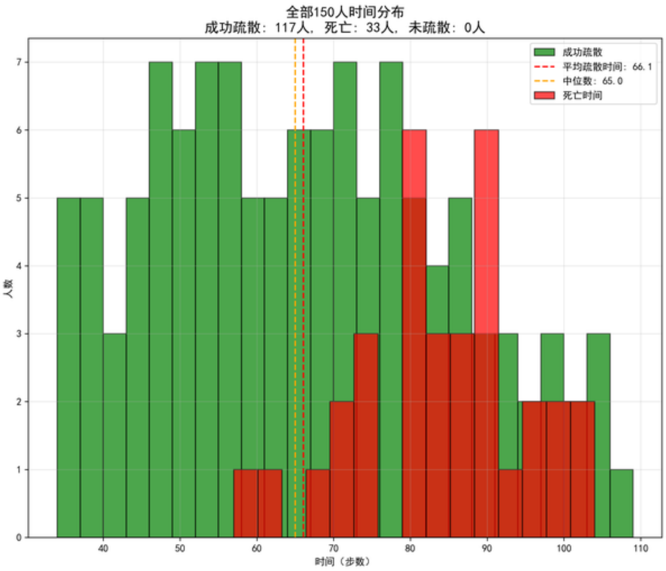
无机器人：

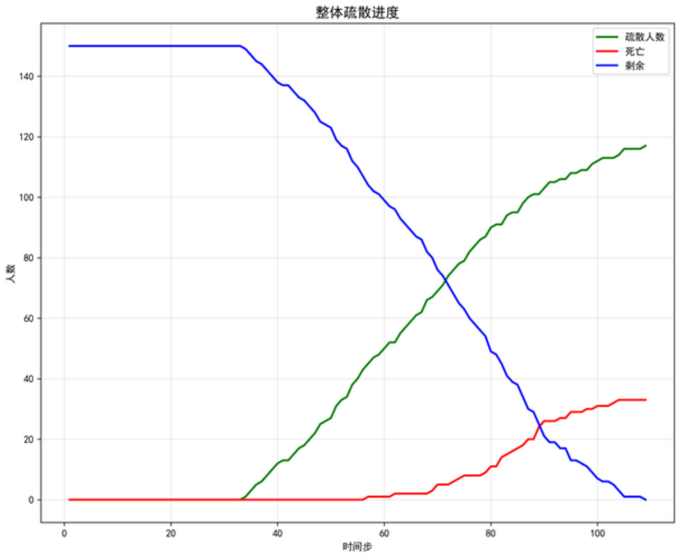


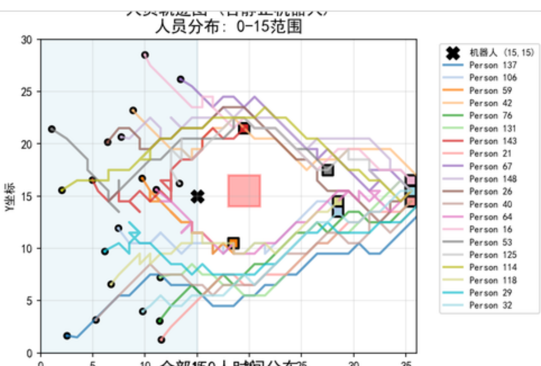

dqn机器人：

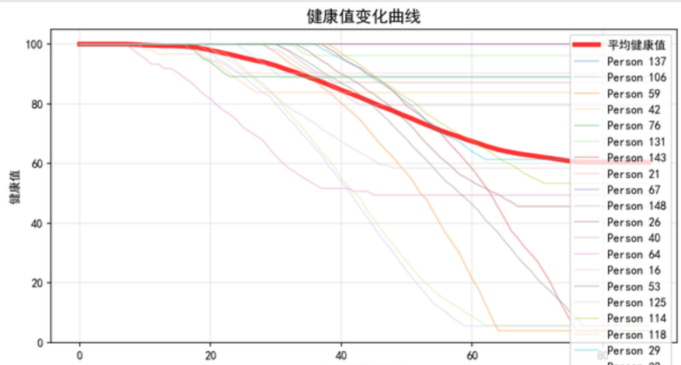

人员轨迹图
人员分布：0-15范围
Y坐标
X坐标
Person 134
Person 79
Person 64
Person 111
Person 8
Person 105
Person 150
Person 58
Person 122
Person 43
Person 48
Person 124
Person 72
Person 22
Person 114
Person 28
Person 5
Person 84
Person 12
Person 21
Robot

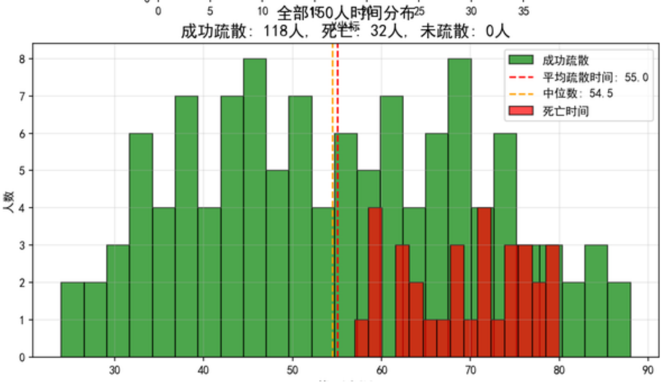

健康值变化曲线
健康值
时间步
平均健康值
Person 134
Person 79
Person 64
Person 111
Person 8
Person 105
Person 150
Person 58
Person 122
Person 43
Person 48
Person 124
Person 72
Person 22
Person 114
Person 28
Person 5
Person 84
Person 12
Person 21

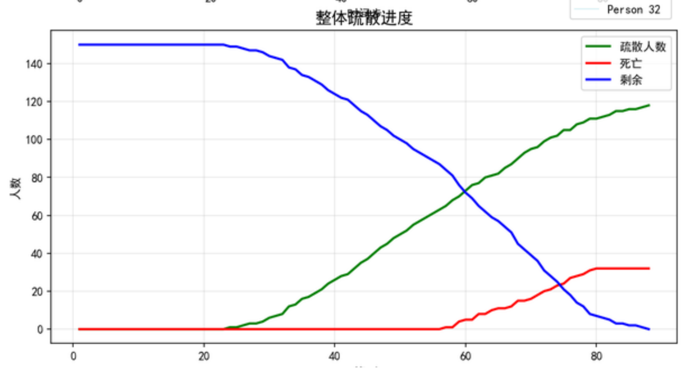

全部150人时间分布
成功疏散：117人，死亡：33人，未疏散：0人
人数
时间（步数）
成功疏散
平均疏散时间：66.1
中位数：65.0
死亡时间

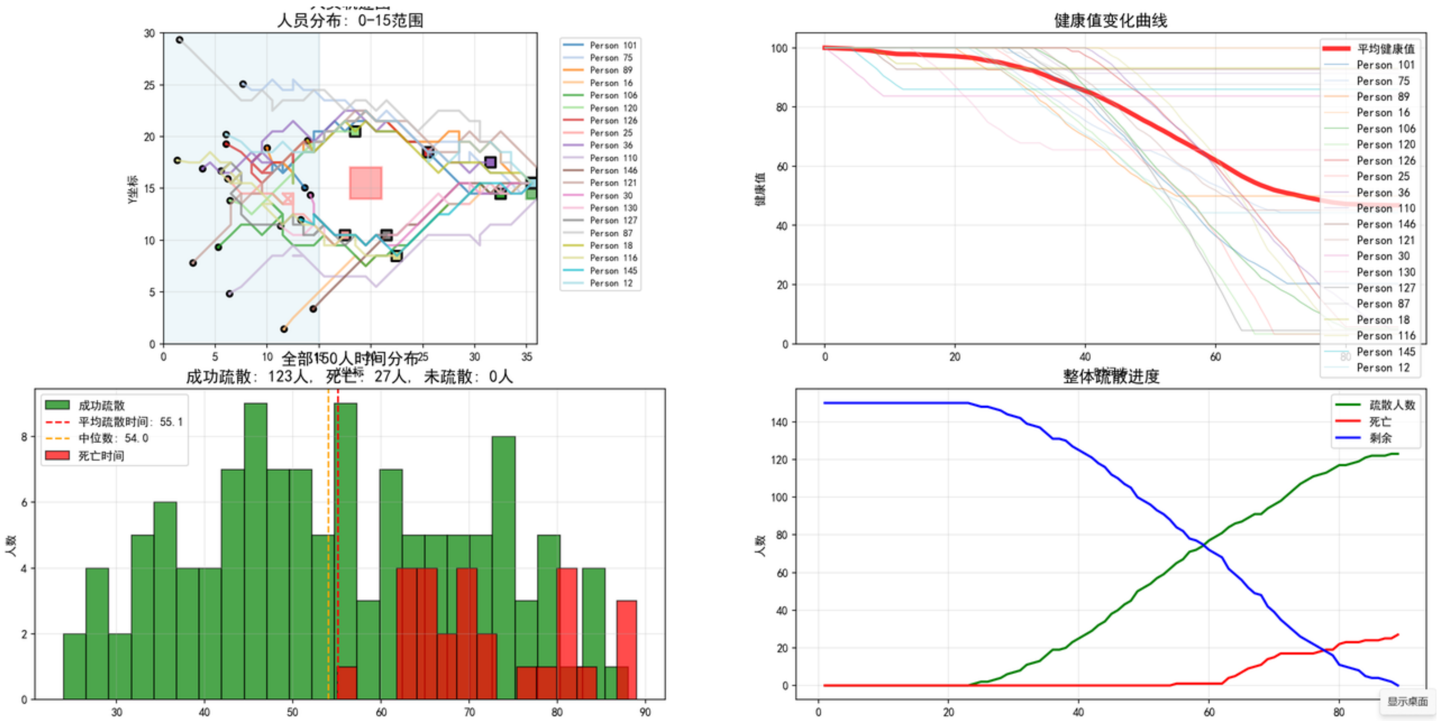

整体疏散进度
人数
时间步
疏散人数
死亡
剩余


人员分布：0-15范围
Y坐标
机器人（15,15）
Person 137
Person 106
Person 59
Person 42
Person 76
Person 131
Person 143
Person 21
Person 67
Person 148
Person 26
Person 40
Person 64
Person 16
Person 53
Person 125
Person 114
Person 118
Person 29
Person 32

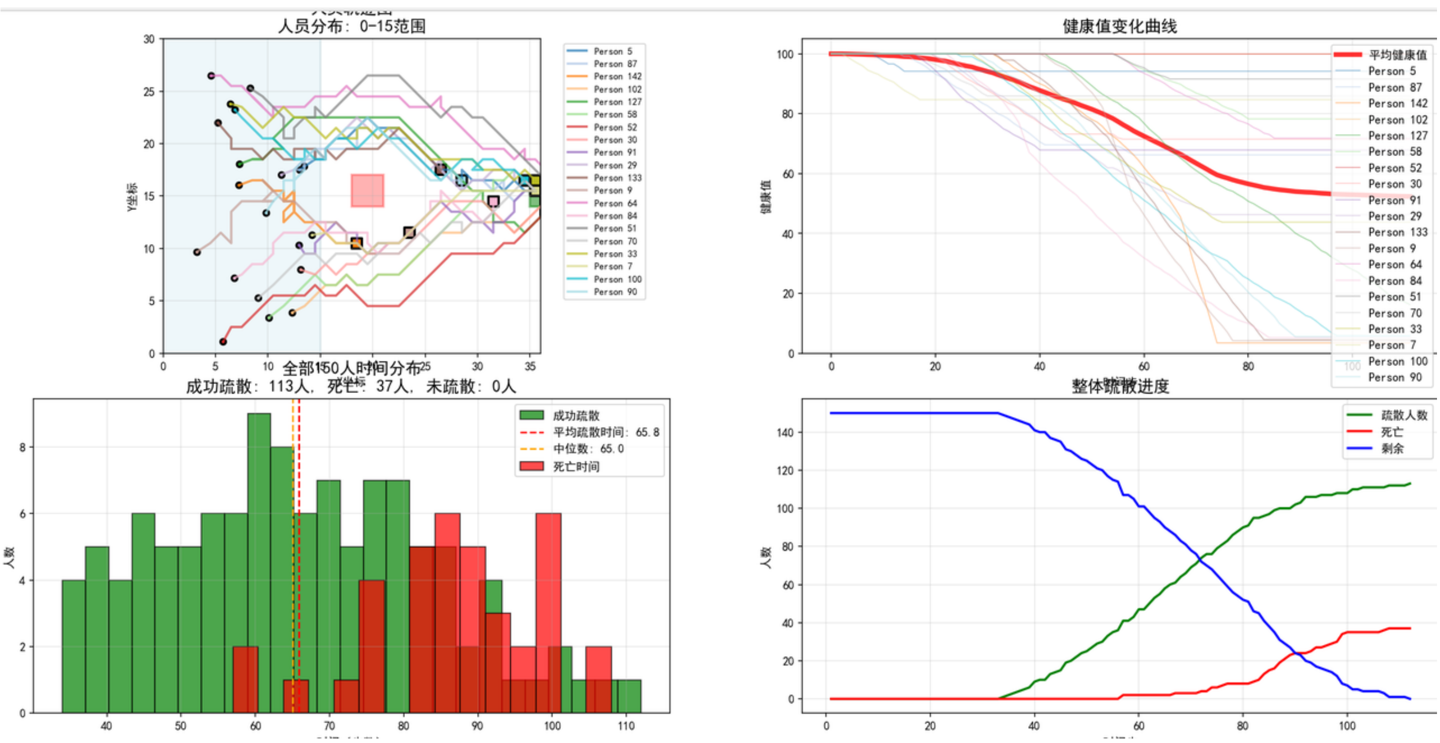

健康值变化曲线
健康值
平均健康值
Person 137
Person 106
Person 59
Person 42
Person 76
Person 131
Person 143
Person 21
Person 67
Person 148
Person 26
Person 40
Person 64
Person 16
Person 53
Person 125
Person 114
Person 118
Person 29
Person 32


全部150人时间分布
成功疏散：118人，死亡：32人，未疏散：0人
人数
成功疏散
平均疏散时间：55.0
中位数：54.5
死亡时间


整体疏散进度
人数
疏散人数
死亡
剩余




无机器人：时间：55.9s

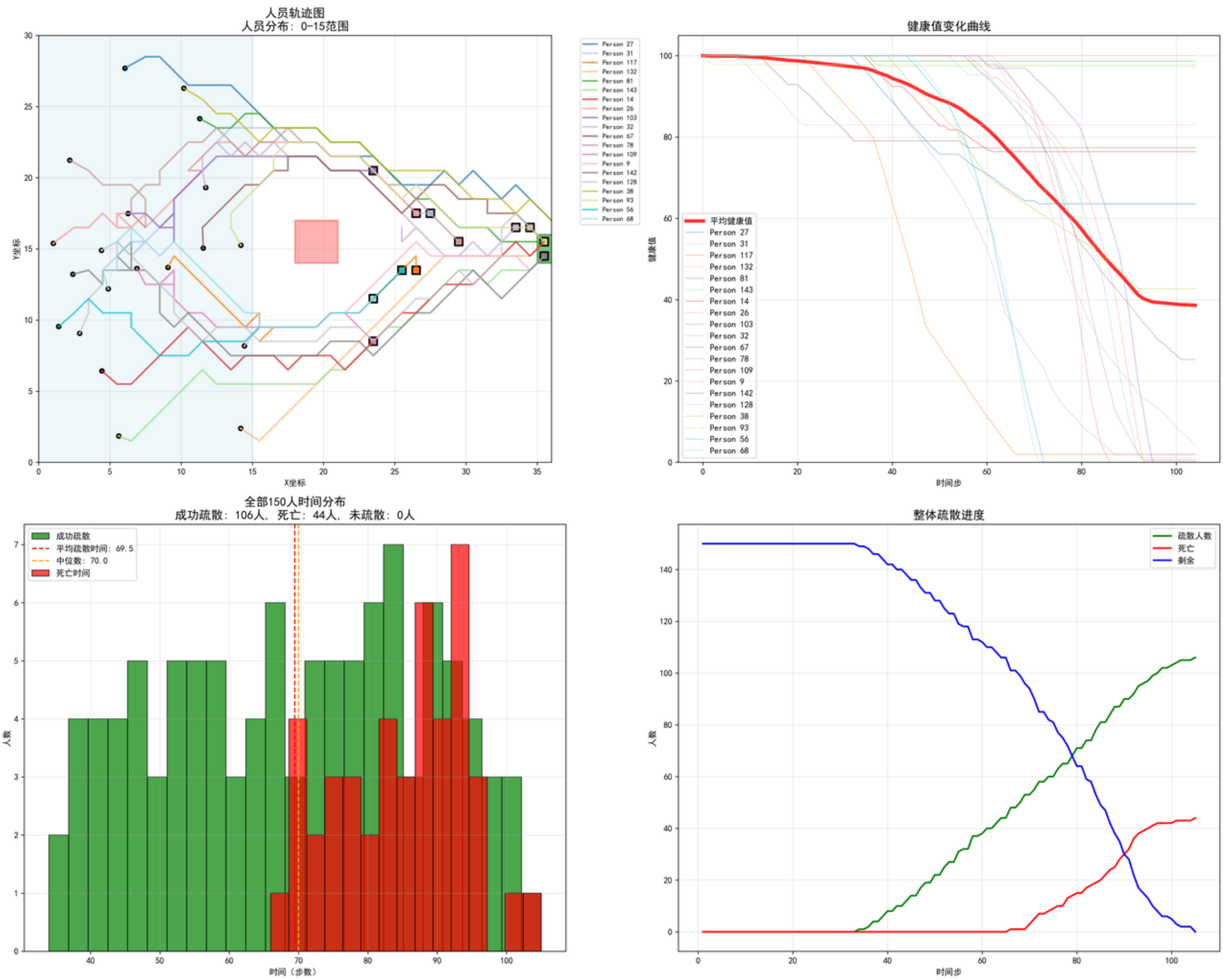



8.4汇报

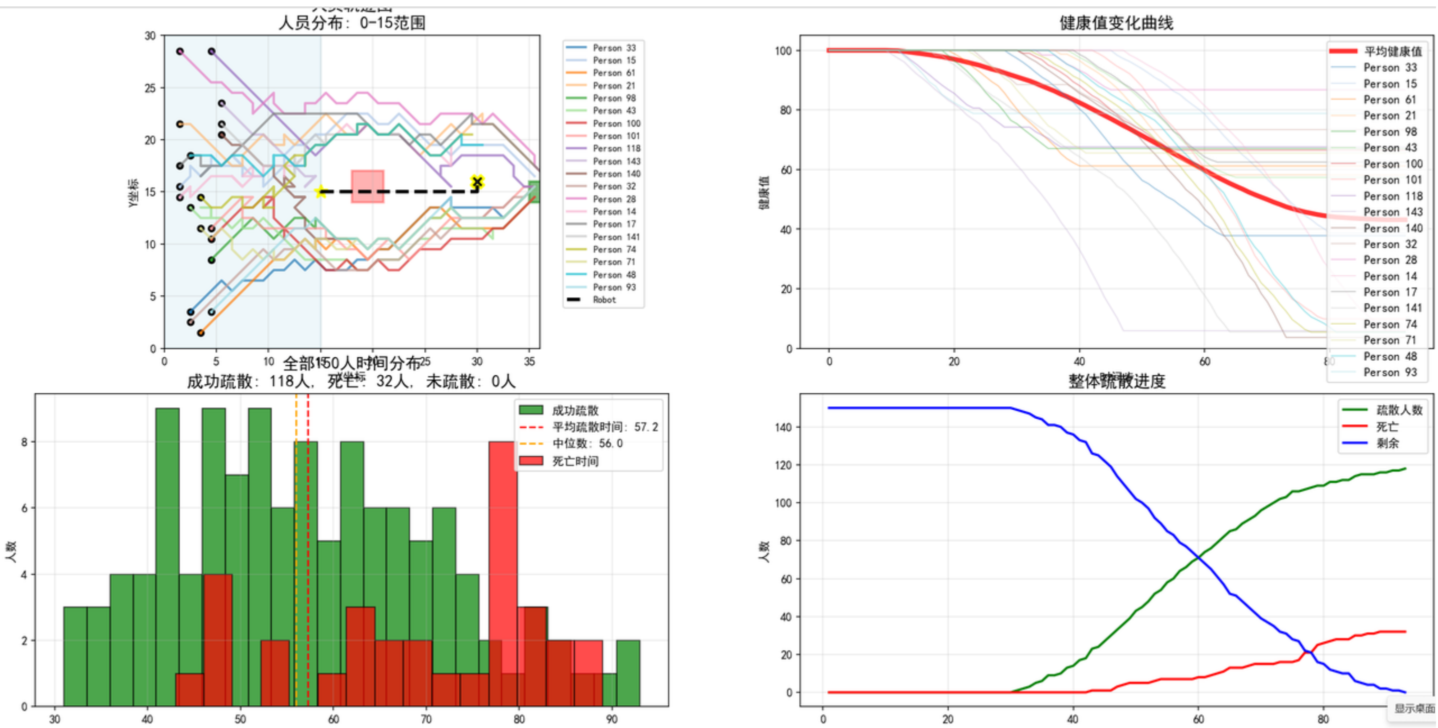
1.奖励函数设计：以时间为主，对于健康值的效果可以相对弱化一点，这样就可以达成想要的效果

2.

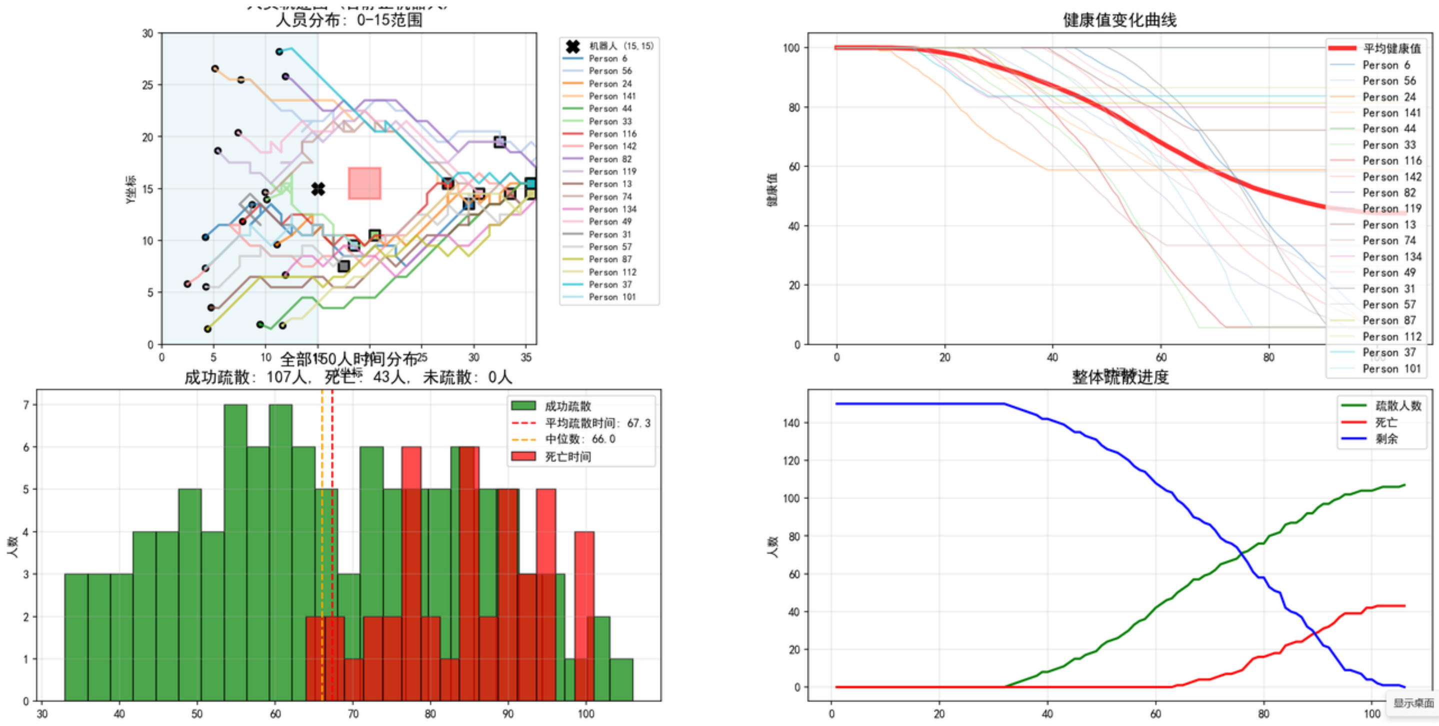
无机器人：时间50.9s




有dqn机器人时间：47.15s(提升约8%)




静态机器人时间 53.5s




假如没有死亡，奖励函数全部变成时间（Chinese）加入健康值后,根据实验结果。加入健康值可能会出现的问题是对于其惩罚机制是多方位的，很容易出现效果不好的情况，Agent的正向反馈可能会与想象不符合，原论文考虑距火平均位置也是这个问题。

### 8.11汇报

#### 1.找文献

shorter evacuation path. In this study, it is assumed that when passing over obstacles, the pedestrians speed reduces to half of his/her original speed.

| 名称 | ISSN | 2025年影响因子 (2025年6月18日公布) 与上一年的差值 | 2025年中科院期刊分区 (2025年3月20日公布) | |
|---|---|---|---|---|
| | | | 大类 | 小类 |
| Automation in Construction<br>Autom. Constr. | 0926-5805 | 11.5 ↑1.9 | 1区 工程技术 | 1区 结构与建筑技术<br>1区 工程：土木 |

#### 2.其他case

机器人在（15,15）处开始，时间：47.15s(时间提升约8%) 最后停止位置（30,15）

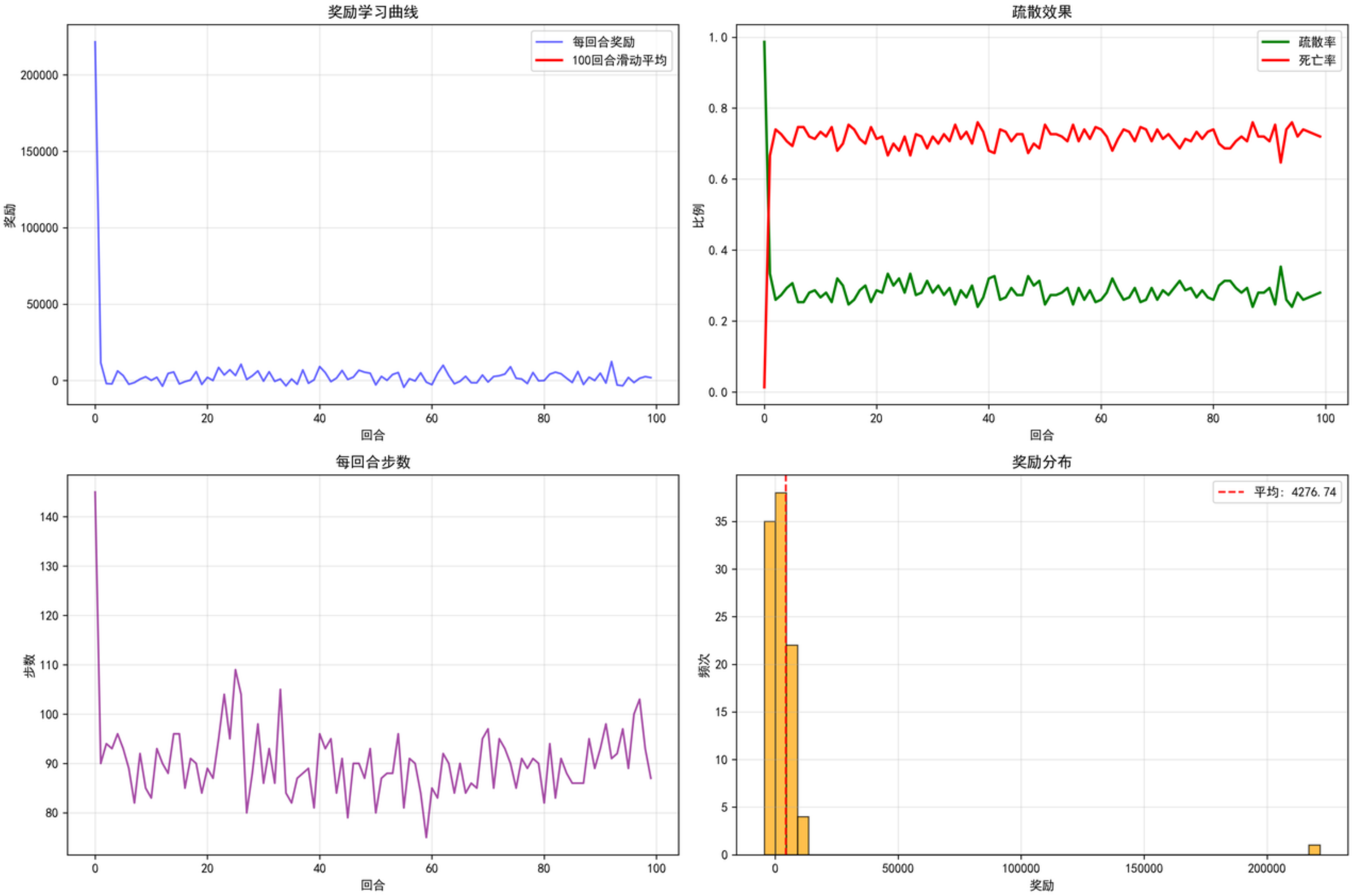
机器人在（30,15）处开始：几乎无移动 时间提升约5% 健康值几乎相同

（师兄提出可以将静止动作移除，小范围内调整）

机器人在（15,20）处开始，健康值甚至会减小（时间提升约9%）最后静止位置大致为（30,15）

**元胞自动机&社会力模型，时间的统一问题**

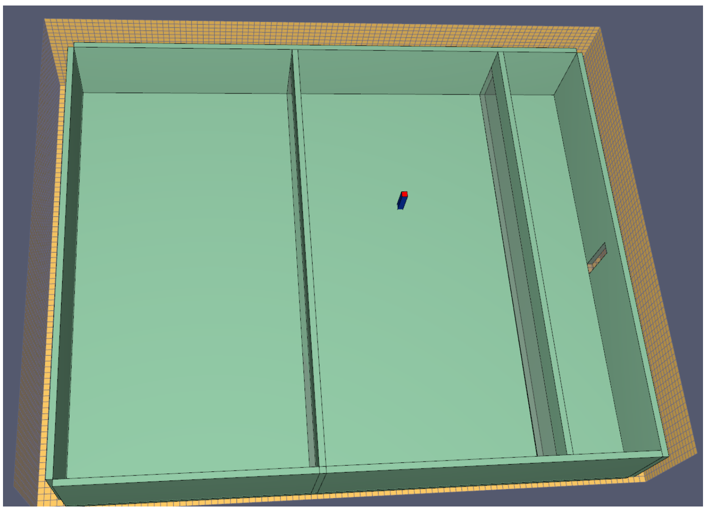
### 3.使用Q_mix算法问题所在：

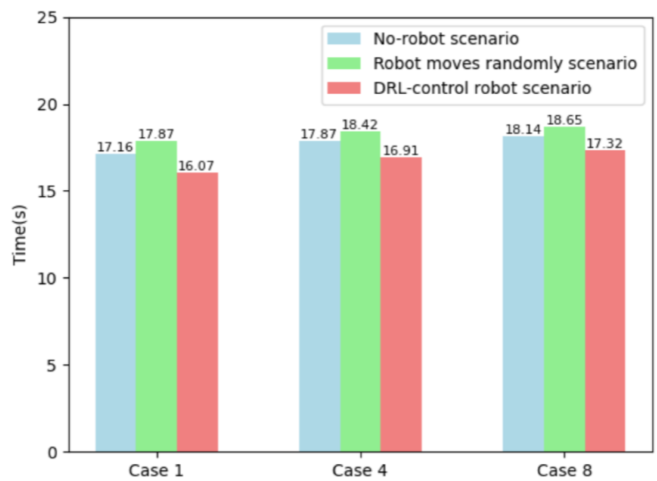



8.18汇报

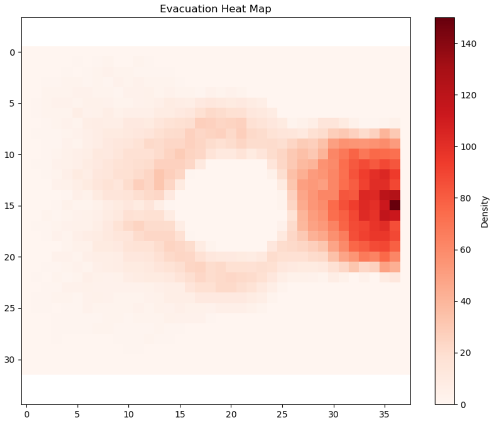
1.把火源的初始大小在代码里面改变了，证明这个确实不太好调

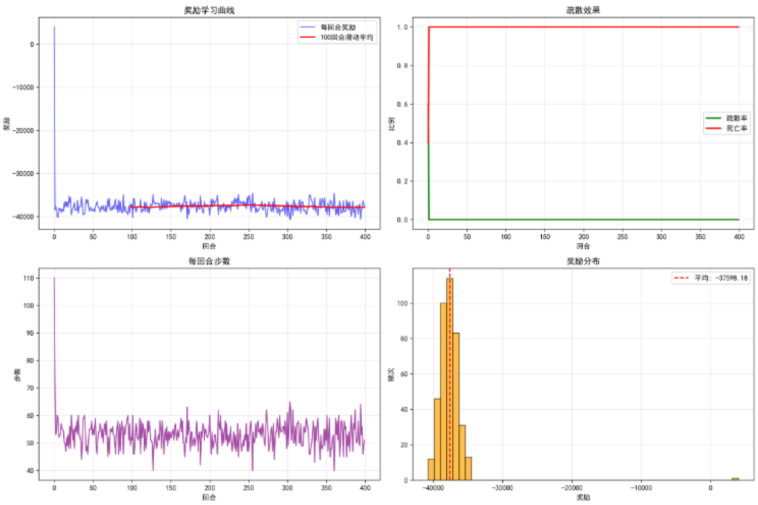
2.改变火的初始位置基本跟Chinese那篇论文结论一致，总疏散时间变少，是否用DQN的差距稍微变少在5%左右

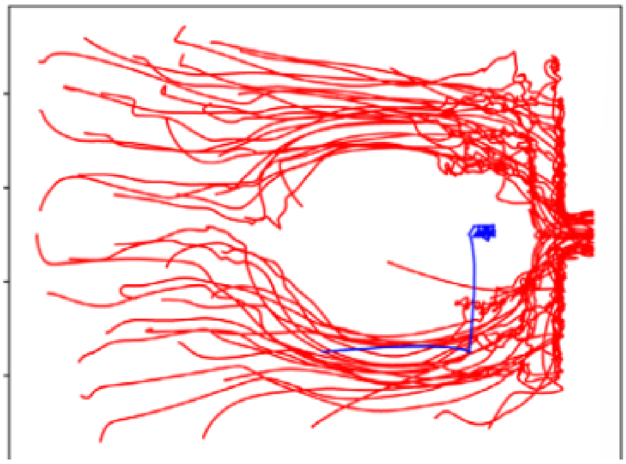




3.Q-mix我基于了上次发的论文里感觉还是不太行，两个机器人排斥效果太强烈了，死亡率就会很高，也可能是场景比较小，关键运行位置较少的原因。这个感觉还是要多调一下

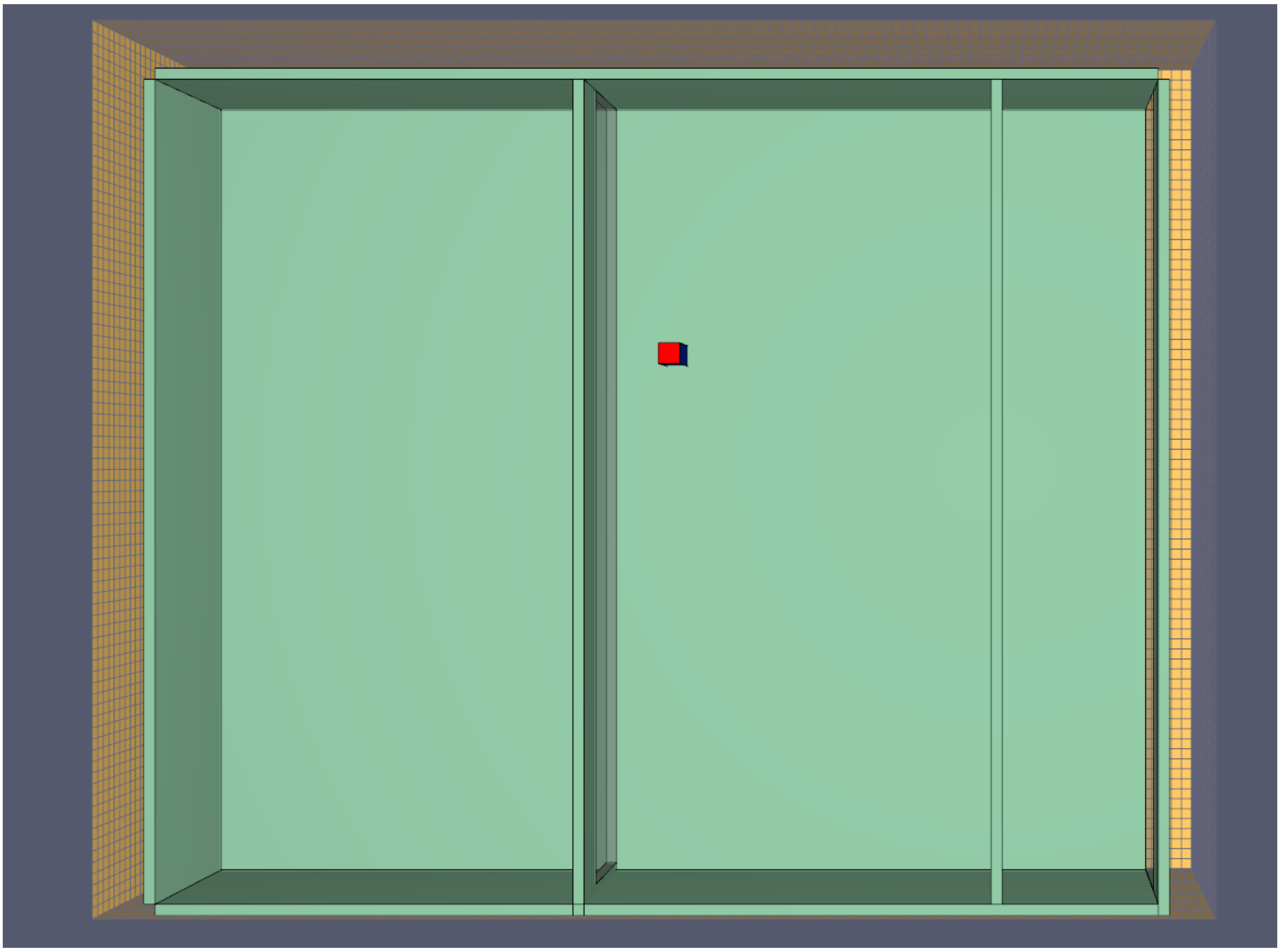


8.25汇报

1.将火源变成两倍



2.qmix

（1）两个机器人若一个机器人落定到最好的位置静止不动，则第二个机器人会起到负面效果，疏散时间会在一定程度降低

经过观察，是两个机器人占用中间过多的地方导致人群绕路，从而使得疏散时间反而增加。（尝试减少机器人的排斥力会效果回升）

（2）动作屏蔽

1.空间区域划分

使用 Voronoi 图算法来自动划分空间。每个机器人占据一个Voronoi单元，疏散进入这个单元的人群

如果采用Voronoi分区，机器人只计算自己所在Voronoi单元内的人群产生的排斥力，完全屏蔽单元外的人群力。

对朝向出口的吸引力进行矢量求和，得到一个期望的运动方向。

问题：人群密度可能高度不均匀实例：两个机器人初始点都位于人少的区域，他们的责任范围覆盖大片空区，而远处的人多的区域无机器人，直到机器人移动过去；Voronoi单元的边界是动态变化的，当人群位于单元边界附近时，两个机器人都认为对方负责，就有责任盲区，没有人会主动处理这些人。

2.定义一组简单的优先级规则，屏蔽低优先级的机器人动作。

*距离出口更近的机器人有更高优先级。它更接近目标，其行动可能更有效。远处的机器人应避免行动干扰到前方机器人的工作。*

*正在排斥一群人的机器人有更高优先级。如果一个机器人已经在对一群人施加排斥力并取得效果，另一个后来者应屏蔽自己朝向这群人的行动，转而寻找其他目标。*

**9.1**

目标：左侧机器人通过先抵达火源伤害处，疏散人群，避免人因过多的遭受到火源伤害

实际：1.机器人速度与人的速度相似（论文略慢），机器人在原文是可穿过火源走直线导致速度降低

2.机器人会向四面疏散人群，刚开始位置选取不对可能会对人员影响大，走的远从而降低速度

3.机器人位置选取还是占绝对优势

通过简单赋予机器人周围的人编号后，疏散平均时间与无机器人进行对比加权得到得分（仅用于观察）

```
(10, 28) -> (11, 29), 可选移动: 6, 最佳得分: -2.77
(15, 32) -> (15, 32), 可选移动: 6, 最佳得分: 22.66
```

存在问题是

1.火源强度和机器人耽误时间没有调节好

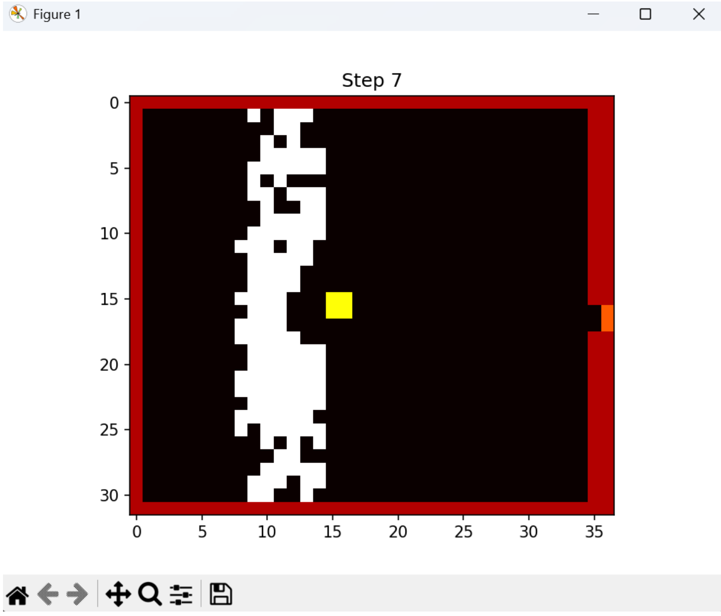
2.机器人速度是否可以稍微比人员快

目前调节1.火源大小和机器人排斥力之间的关系（qmix调好一点，带入dqn会相应的差一点）2.机器人的初始位置

**9.9**

完成更细致机器人的位置范围确定

调节机器人的初始位置




目前任务：1.在pyrosim改变火源的位置和人员数量是否变化（52.25s）

2.调节速度与健康值的关系，尝试是否可以达到第二个机器人效果更好的状态

3.按照论文可以实现0.05s运动0.05m

左侧机器人初始位置不同（秒是按帧除的）：

（15,15） -> (14,15) 47.25s

（15,10） -> (14,15) 48.25s

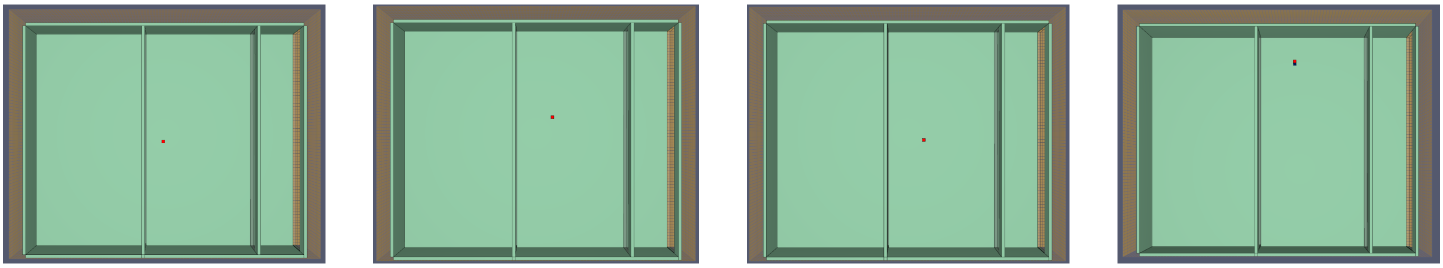
（10,10） -> (14,15) 50.25s

（10,15） -> (14,15) 49.75s

（13,15） -> (14,15) 47.25s

（13,10） -> (14,15) 48.75s

**9.29**



a　　b　　c　　d(效果欠佳)

首先在不同的场景下对于火源放置不同

对于疏散时间和健康值以及最后的疏散人数详情图：

无机器人：疏散时间：50.9s　　　平均健康值：72.95

开始设置的是全局，但是试验过c情况后，考虑到当火源放置远后可能会出现健康值急剧增高

a.左侧机器人初始位置不同（秒是按帧除的）：

（15,15） -> (14,15) 47.25s （8%）　平均健康值：77.17

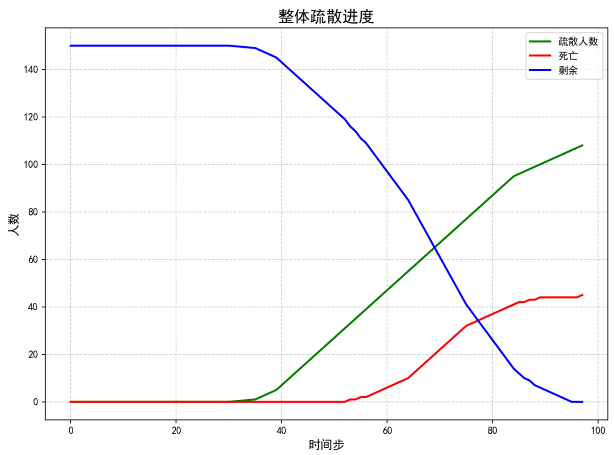
（15,10） -> (14,15) 48.25s　　　平均健康值：76.23

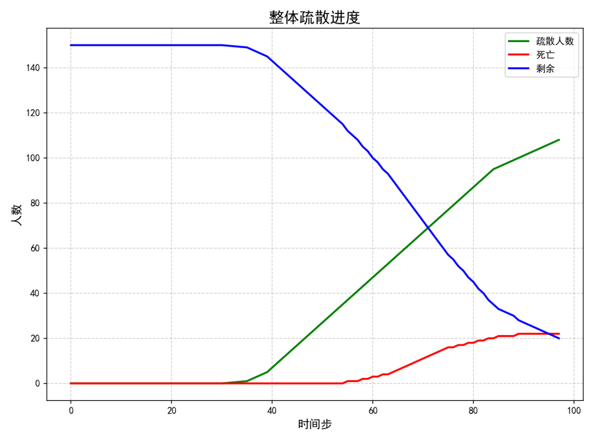
（10, 5） -> (14,15) 49.25s　　　平均健康值：76.16

（10,15） -> (14,15) 49.75s　　　平均健康值：76.51

（13,15） -> (14,15) 47.25s　　　平均健康值：76.27

（13,10） -> (14,15) 48.75s　　　平均健康值：76.81




b.左侧机器人初始位置不同：

（15,15） -> (14,18) 47.25s　　　平均健康值：78.15

（15,10） -> (14,18) 48.75s　　　平均健康值：77.18

（10,10） -> (14,18) 50.25s　　　平均健康值：77.87

（10,15） -> (14,18) 49.75s　　　平均健康值：77.26

（13,15） -> (14,18) 47.25s　　　平均健康值：77.19

（13,10） -> (14,18) 48.75s　　　平均健康值：77.01

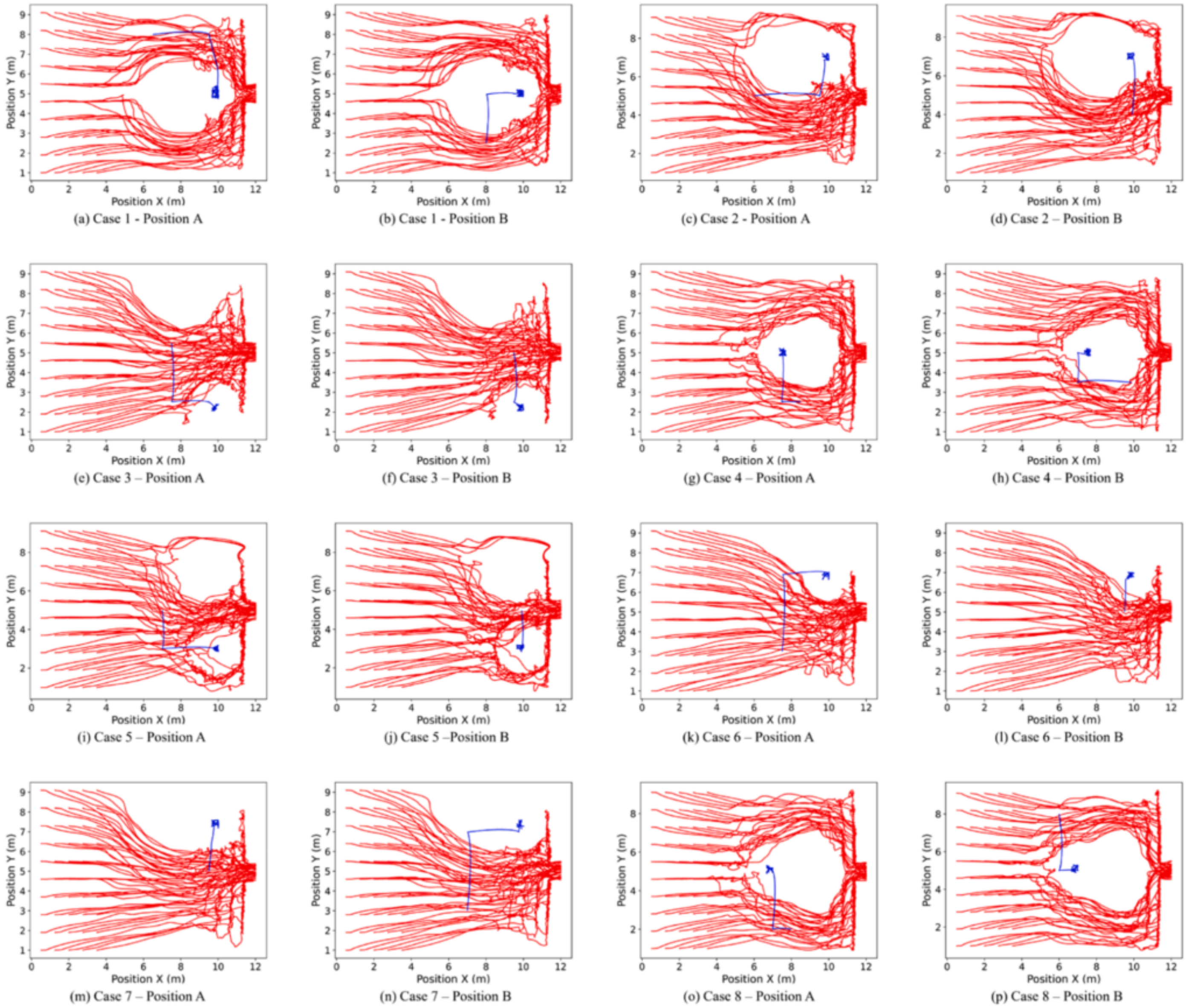
d.

无机器人

（15,15） -> (14,18) 49.75s　　　平均健康值：70.13

（15,15） -> (14,18) 49.25s　　　平均健康值：84.21

目前想法：剔除掉健康值为一个特定值以上的人群。




As illustrated in Fig. 14(a,b,d,e and h), for Case 1, Case 2, Case 4, Case 5, and Case 6, robots within a specific area exhibit a slight increase of 0.1-0.3 meters in the average distance between pedestrians and the fire hazard area during the middle and late stages of evacuation. This observation suggests that, in these cases, robots under DRL control do not adversely impact pedestrian safety. Fig. 14

论文中：行人与火源的距离增加约 0.1–0.3 米

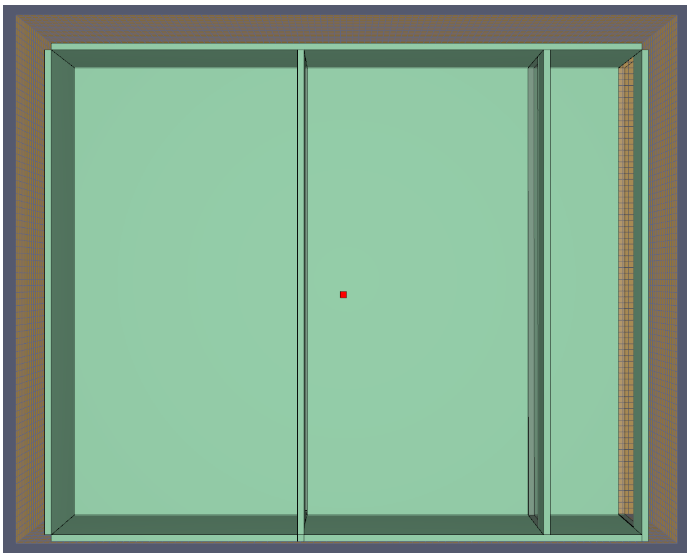
**10.13**

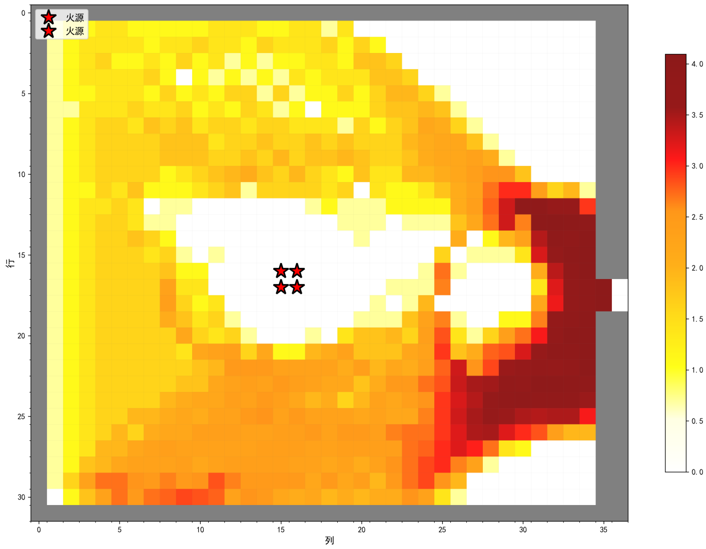
1.行人轨迹图以及相关的热力图

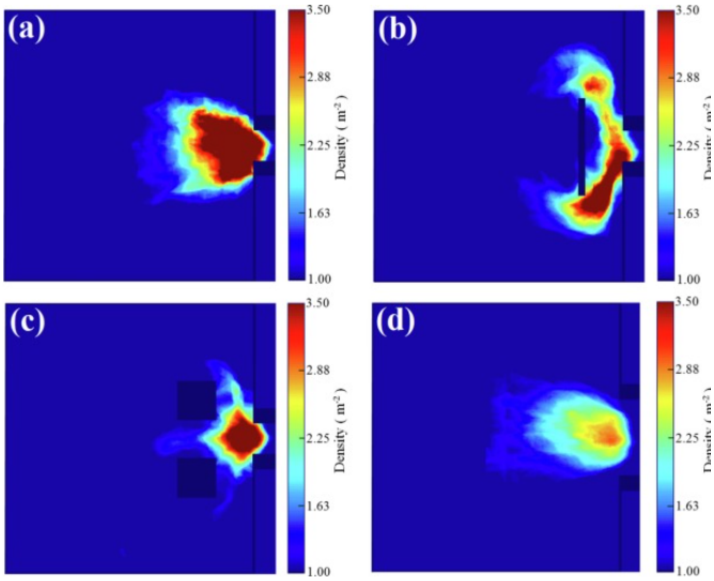
a 机器人1从（10,10）出发效果更好

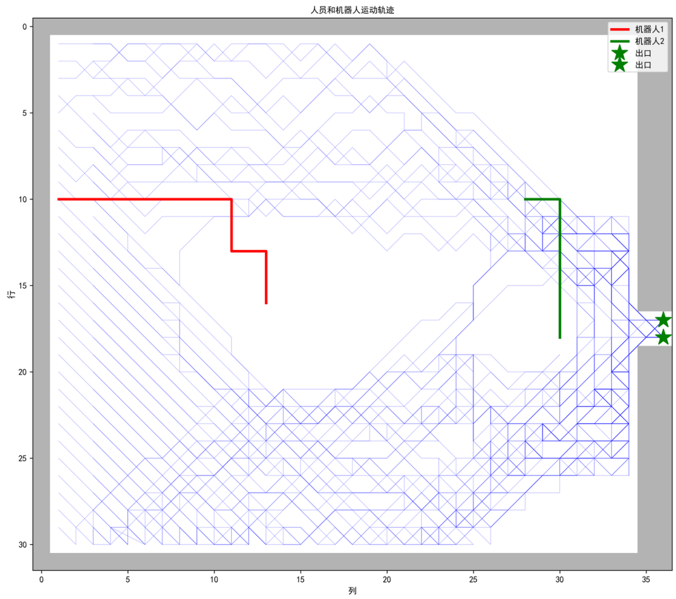
放置的是机器人1 从（10,1）出发的效果图，左下是调整思路（与长条状障碍物类似故而能较好地提高疏散效率）

因为元胞自动机的缘故，中图很难表现出如社会力模型中每个人员人员清晰的路线。



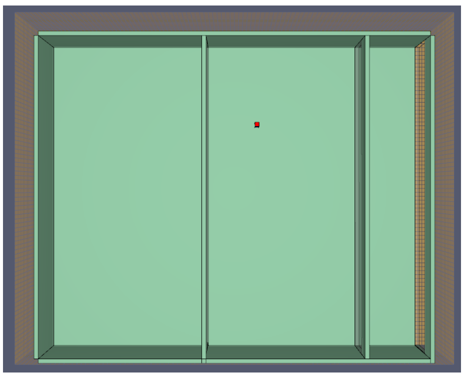


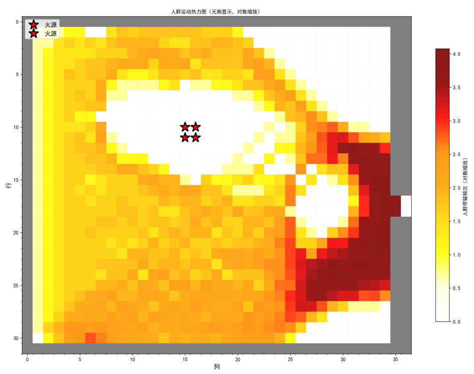



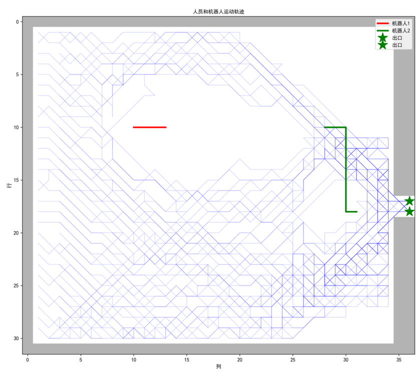




（10,10） -> (14,18) 47.75s　　　平均健康值：75.87

b







（10，1） -> (14,15) 49.25s　　　平均健康值：76.16

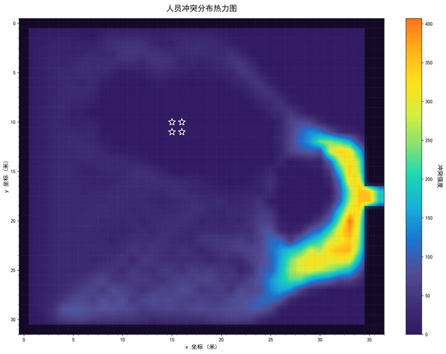
2.若火源较远机器人的运动位置

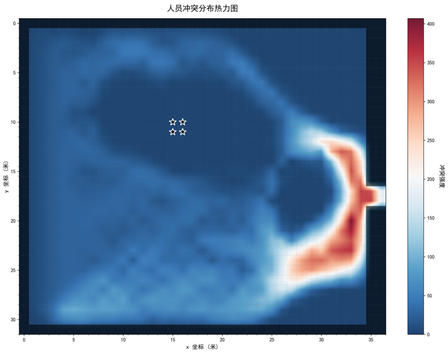
机器人会直接向上向下走避免影响人的运动。（可能做一个补充）

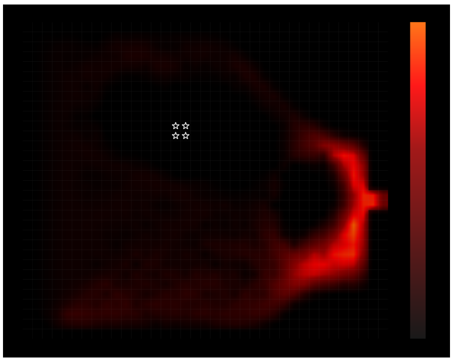
10.23

用PDF处理后的图像（以火焰位置二为例）：

右侧更能表现机器人的影响范围，中图与论文图更为接近。







以不同人数为标准：

人员疏散出80%记作完成疏散

以下选取最有代表性的机器人**初始位置**、火焰位置一的数据：（不同火焰位置的大致结果趋势一致）

人数：50,100,150,200,~~250,300,350,400~~

均匀分布对于人数较少的影响较大

**50人：**无：25.0s有：25.5s　均匀分布误差较大，几乎没有形成成拱现象

**100人：**无：36.25s 有： 34.75s

**150人：**（现有的）无：49.5s 有：46.25s

**200人：** 无：62.0s 有：59.75s　　机器人效果良好，由于调试火源大小及火灾影响没有

**250人：**无：75.25s 有：72.5s 失能影响结果，该结果是不考虑失能。250及以上由于火源扩散，存在人员健康值过低，疏散时间约等于死亡时间，没有比较意义

~~**300人：**~~无：87.75s　　不加入机器人疏散效果更好

~~**350人：**~~无：100.5s

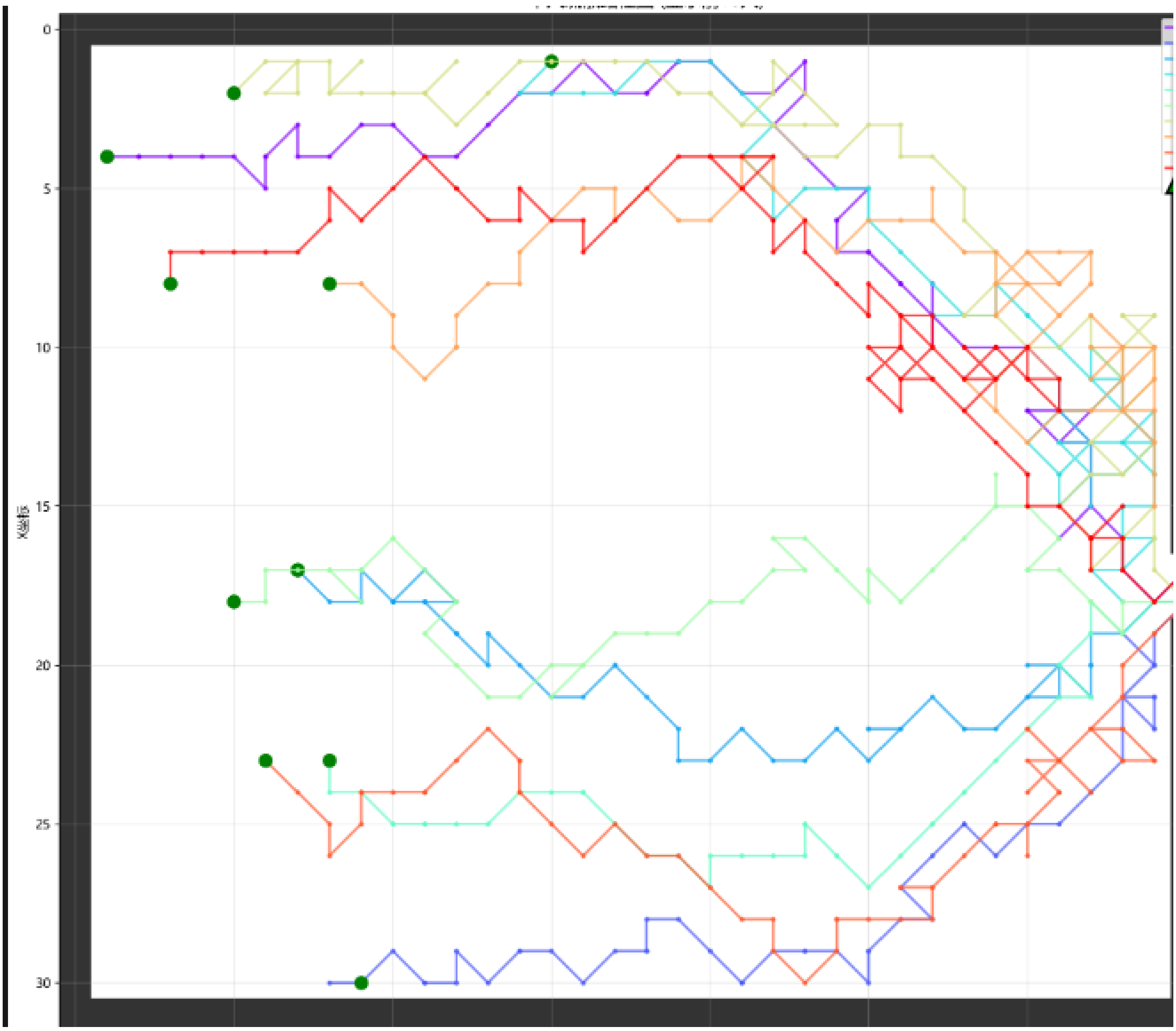
~~**400人：**~~ 无：113.25s

11.3

1、关于前面当人数过多的时候

强化学习效果差是只在这个有机器人的情况下比机器人到一边效果差

例如：

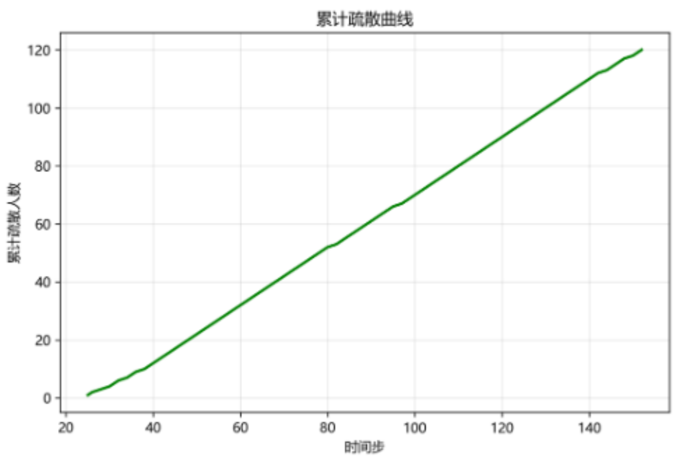


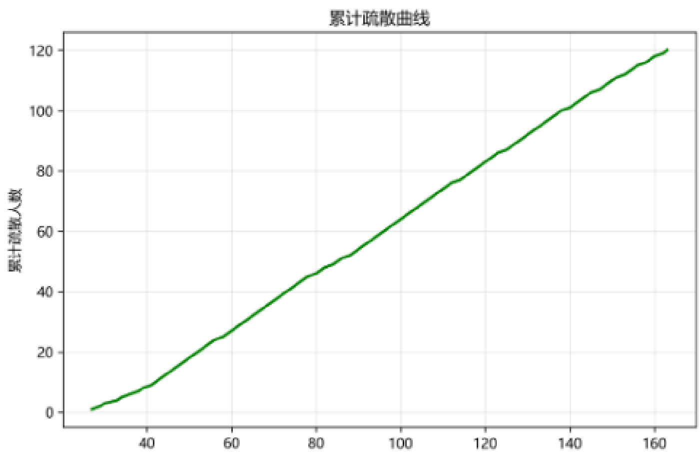

这两种情况都比没有机器人的效果要差。

人员疏散图：使用方法：降低火源强度，增加机器人的排斥强度等方法

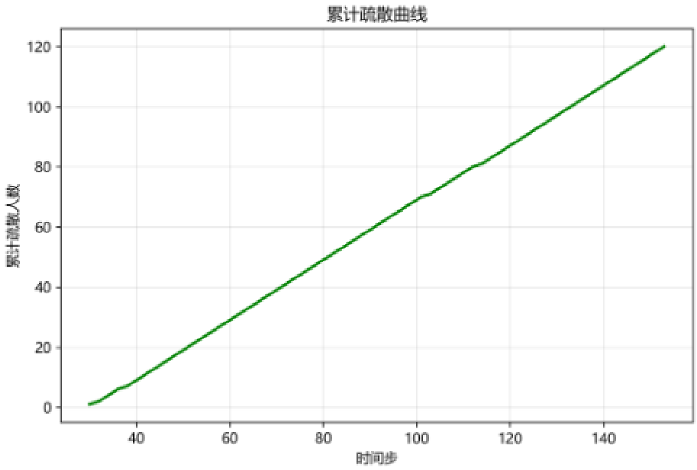
150人 左侧有机器人，右侧无机器人（7.24%）

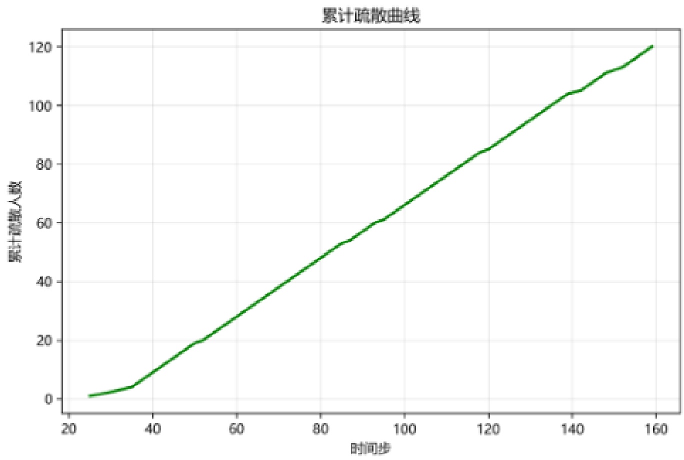
152 163



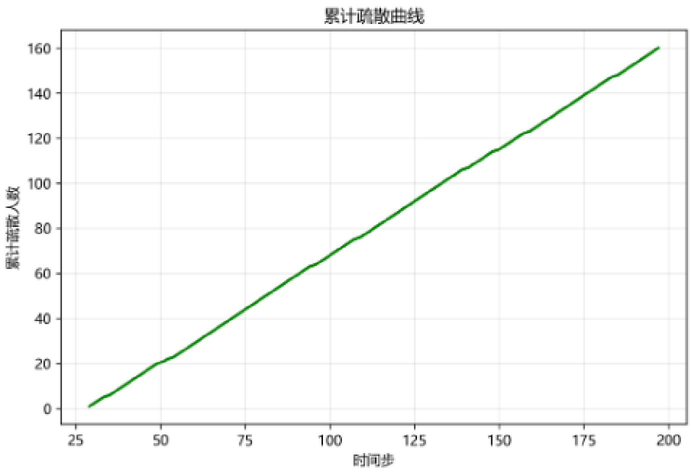

降低火源强度后出现疏散均有所提升但是总体比例没有变得更好

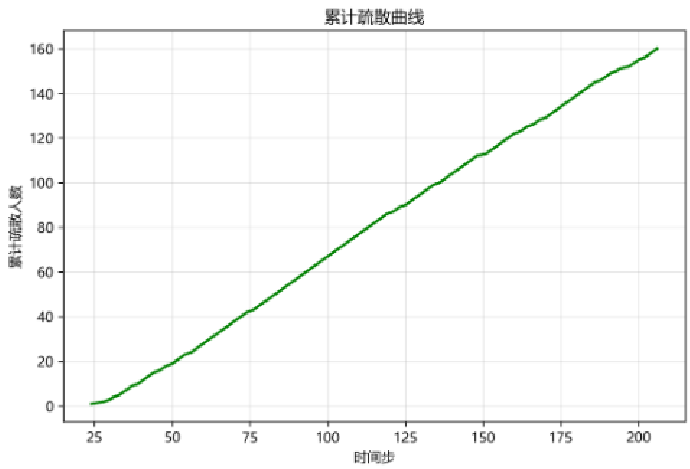




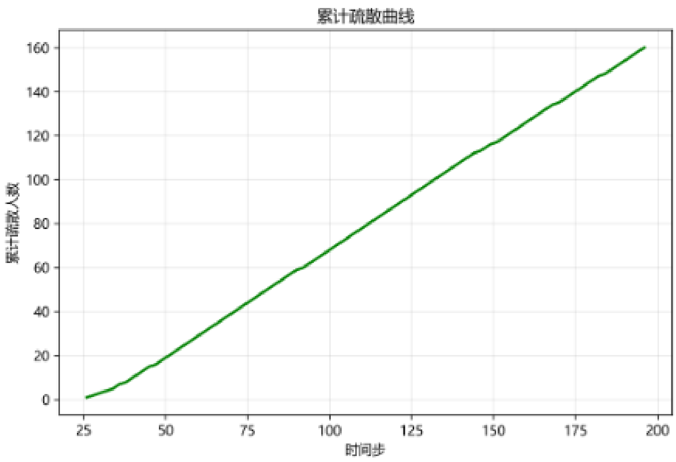

200人 左侧有机器人，右边无（在左侧初始时间比右侧慢的情况下更先疏散完）（4.37%）

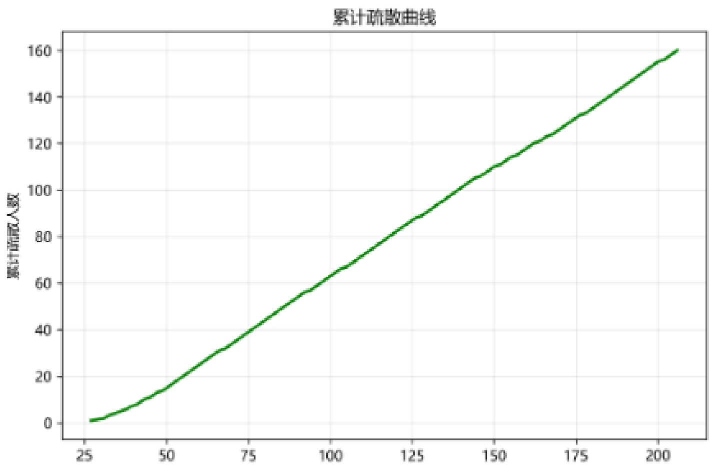
197 206





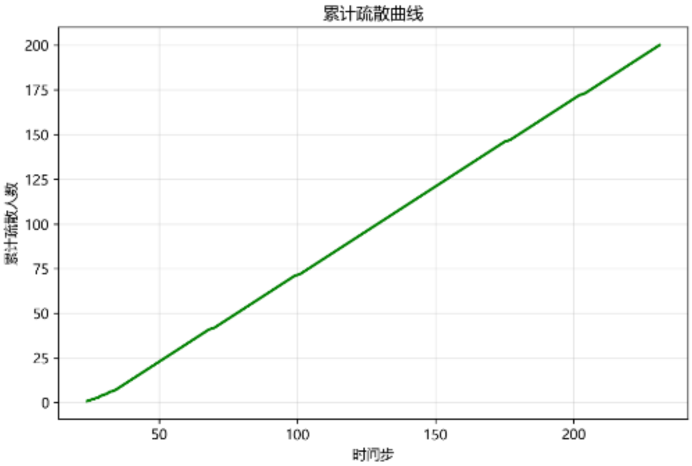

196

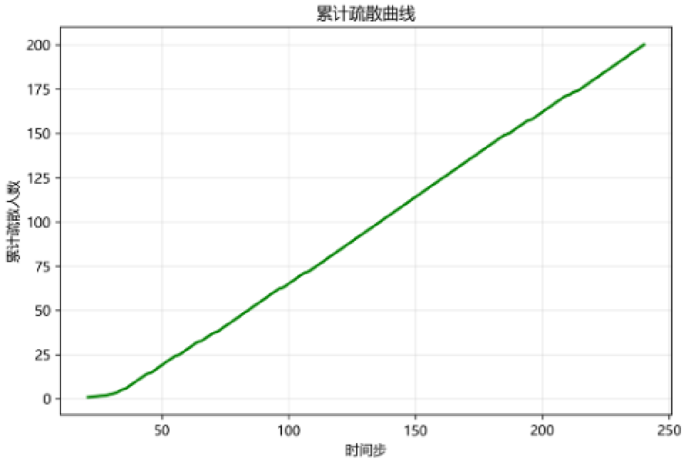


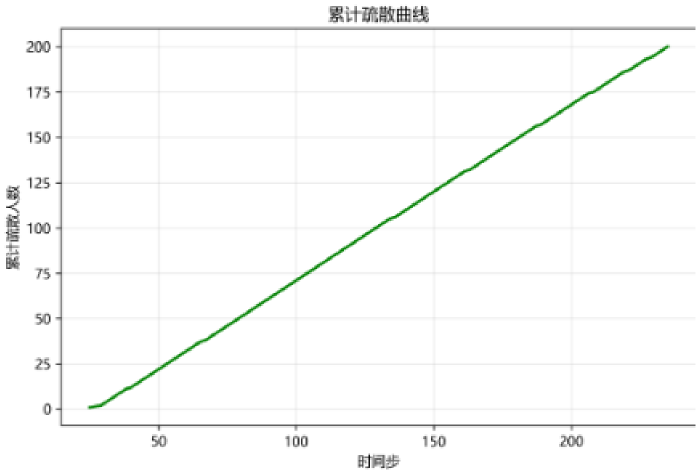


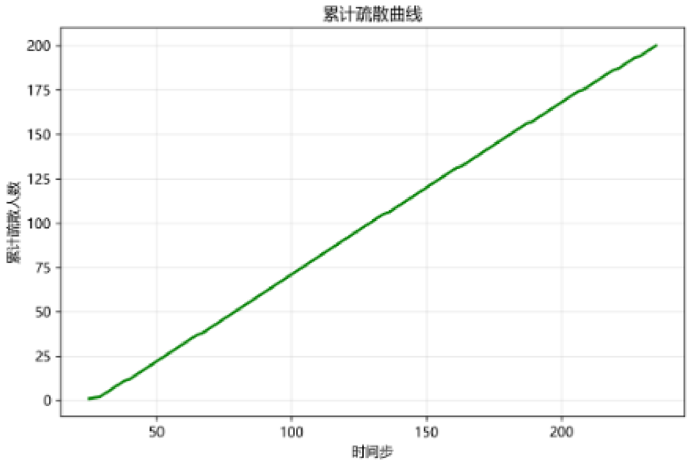

250人 左侧机器人，右侧无机器人（3.75%）

231 240


累计疏散曲线
累计疏散人数
时间步


累计疏散曲线
累计疏散人数
时间步


累计疏散曲线
累计疏散人数
时间步


累计疏散曲线
累计疏散人数
时间步